PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 16 March 15, 1930 Number 3

CIRCLE FORMATION IN DATURA AND OENOTHERA

By Albert F. Blakeslee and Ralph E. Cleland[1]

Carnegie Institution of Washington, Department of Genetics, Cold Spring Harbor, N. Y., and Goucher College, Baltimore, Md.

Communicated February 13, 1930

At the Science Meeting in New York, December, 1928, one of us (C) presented a summary of the occurrence of circles of attached chromosomes which he had discovered in hybrids between complexes in Oenothera[2] and the other (B) summarized the morphological effects and configurations induced by extra chromosomes in Datura and the circles induced in hybrids between cryptic types in nature and a single line used as a standard.[3] The chromosomal phenomena in Oenothera and Datura appeared to parallel each other in so many respects that it seemed desirable to discover, if possible, in how far the two forms had a common basis for their chromosomal behavior. As a result of conferences, a joint paper[4] was presented before the American Philosophical Society last April, which is the basis of the present and immediately succeeding articles.

Students of Oenothera have led the way in the discovery of chromosomal duplication as the cause of mutant types such as tetraploids, triploids, and $(2n + 1)$ forms. They have also led in the discovery[5] of circles of 4 or more attached chromosomes as a regular phenomenon in species and hybrids. The discovery that heterozygous species with permanent circles are made up each of two complexes responsible for "twin hybrids" or unlike reciprocals when they are crossed with other types is due largely to the investigations of Renner.[6]

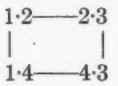
In Datura a duplication series has been discovered similar to that found in Oenothera but lethals are not normally present and the chromosomal phenomena in general resemble more nearly those in other plants. For these and other reasons, especially the fact that the chromosomes of Datura can be distinguished cytologically or located in small size classes, an analysis of the chromosomal and breeding behavior has been a simpler task than in Oenothera. In Datura it has been possible to distinguish among the $(2n + 1)$ types: (1) *Primaries* in which the extra chromosome is like the other two in the trisome, (2) *Secondaries* in which the extra is

made up of one half of a normal chromosome doubled by interchange between homologous chromosomes, and (3) *Tertiaries* in which the extra is made up of parts of two non-homologous chromosomes combined by segmental interchange.[3] From the peculiarities of these extra chromosomal types, considerable information has been gained as to the factors in parts of chromosomes and the unbalancing morphological effects to be expected when such parts are present in excess. Such forms, therefore, have been found useful as testers in analysis of new chromosomal types. By means of trisomic ratios of genes it has been possible further to label chromosomes and parts of chromosomes and to follow their breeding behavior in Primaries, Secondaries and Tertiaries.

In Primaries the 3 chromosomes of the trisome are attached though never in a closed system, while those of Secondaries can form a closed circle of 3. In Tertiaries that have been analysed, the extra chromosome may be attached as a connecting link between two bivalents. The chromosomal parts involved in such types have been determined by chromosomal sizes, by chromosomal types produced in their offspring and by inheritance of genes. Triploids, and tetraploids have their chromosomes attached in 3's and 4's, respectively.

Cryptic types have been discovered in nature as well as induced by radiation treatment, which when crossed with a standard line cause in meiosis the appearance of closed circles or chains of 4 to 6 attached chromosomes.

Among these, the "B" type has been most thoroughly analysed. The circle of 4 induced by crossing "B" races with a standard line seems best explained on Belling's hypothesis[7] of the origin of such races by segmental interchange between non-homologous chromosomes. The hypothesis was proposed to account for the attraction between chromosomes of unlike size in a $(2n + 1)$ type derived from hybrids involving a "B" race. It was predicted, however, that if the hypothesis were correct a circle of 4 should result in diakinesis.[8] The method of interchange has been represented by diagrams[3] but can also be shown by numerals. If 1·2 and 3·4 represent two chromosomes in the standard line and the end 2 should have interchanged with the end 4 in the origin of the "B" race, the formula for the resulting 2 chromosomes in the latter race would be 1·4 and 3·2. In the hybrid between the two races, therefore, attraction of like ends would bring about a circle of 4 chromosomes in diakinesis with the following arrangement:



The two chromosomes in the "B" circle can be distinguished cytologi-

cally from the other 10 chromosomes in Datura by morphological differences. The parts of these chromosomes that have been interchanged has been determined by the morphological peculiarities produced in the plant when either of the two "B" chromosomes is present as an extra. The conclusions have been checked by a number of lines of evidence including transmission of genes, production of chromosomal types in the offspring and types of chromosomal configurations when $(2n + 1)$ types are made heterozygous for "B" races. Breeding behavior, together with the cytological findings of Dr. Bergner, shows that crossing-over takes place between "B" and Line 1 chromosomes of the attached configuration in hybrids between these races.

As has been pointed out elsewhere,[3] the various configurations analyzed in Datura, some of which have been mentioned above, seem explainable only on the assumption of attraction of like ends. On this assumption it has been possible successfully to predict the configurations that would result in hybrids between different chromosomal types. The formation of a circle of 8 chromosomes in hybrids between a "B" tetraploid and a Line 1 tetraploid is such a prediction confirmed by the cytological findings of Dr. Bergner.

In the case of the "B" circle of 4, the evidence, especially from the lack of aborted pollen in such hybrids and from the breeding behavior, has led to the belief that like parts are constantly separated to opposite poles at disjunction. This conception of attraction of like parts preceding disjunction and the separation of like parts at disjunction is but an extention of the normal behavior of homologous whole chromosomes to the behavior of their parts when the latter have been interchanged or translocated. That homologous ends are normally adjacent at pairing is shown in other forms: by chromosomes, the two ends of which are distinguishable by presence of arms, and by the location of chromomeres of different sizes as well as by other evidence. It is also at the basis of the present conception of linkage and crossing-over. It is not yet known, however, in what the likeness of ends consists, which makes them attract. The hypothesis that such attraction brings together homologous parts of chromosomes seems amply justified by all the chromosomal phenomena in Datura so far analyzed.

Håkansson[9] has attempted to apply the hypothesis of segmental interchange to the attachment of chromosomes into the circles in Oenothera but believes he finds insuperable difficulties in the origin and behavior of half-mutants. Darlington,[10] whose article appeared after we had worked out the diagrams of complexes given in the previous paper,[4] has ably defended the hypothesis of segmental interchange for Oenothera and met the objections regarding the half-mutants. It will not be necessary, therefore, to consider here in detail the matters discussed by him.

It will suffice to show that the types of configurations found in combinations between different complexes of Oenothera are not inconsistent with a belief in the attachment of adjacent chromosomes by homologous ends. The present paper will further point out some of the consequences that would be expected if chromosomes in Oenothera have been altered by segmental interchange.

Among the genera of plants which in diakinesis and the heterotypic metaphase exhibit chains or circles of chromosomes attached end to end, the most striking is Oenothera. In this genus, a great majority of forms (at least in the "*biennis* series") have circles which comprise most or all of the chromosomes. Each of such species or hybrids is composed of 2 associated chromosome complexes which differ in gene content, and, at least in spontaneous species, these complexes are passed on intact and unbroken from generation to generation of a selfed line.

There have been discovered to date 12 different chromosome arrangements of the 15 which are theoretically possible in 14-chromosome forms. The number of ways, however, in which the ends of two complexes of 7 chromosomes each can be arranged, in order to form, when these complexes are brought together, any chromosome configuration possible to the genus Oenothera, is so large that this fact in itself is of little significance. It is much more important to discover whether the various configurations found in Oenothera could have arisen gradually in the process of evolution from the unmodified ancestral type which undoubtedly had all of its chromosomes paired. It is easy to see, however, that even the largest circles could have come about by comparatively short series of simple interchange. To take a simple series: Let the arrangement of ends in the chromosomes of the ancestral form be represented by the numerals 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14.

An interchange between the first two chromosomes similar to that believed to have occurred in the origin of "B" races of Datura discussed above would produce a complex with the formula 1·4, 3·2, 5·6, 7·8, 9·10, 11·12, 13·14. A hybrid between this complex and the original form would show a circle of 4 and 5 pairs. Should an exchange of ends occur between one of the modified chromosomes and an unmodified chromosome, a new complex would be formed, the formula of which may be represented as 1·4, 3·6, 5·2, 7·8, 9·10, 11·12, 13·14. This combined with the original complex would give a circle of 6 and 4 pairs.

In similar manner, complexes might be formed capable of giving circles of 8, 10, 12 or 14 with the original complex. Granted the early appearance of balanced lethals in the chromosomes of the circle, and consequent true-breeding for the circles, strains might conceivably have arisen in the course of evolution containing not merely these but other types of configuration as the result of comparatively short series of simple interchanges.

The fact that circles are large in many species of Oenothera does not then constitute a serious obstacle to the segmental interchange hypothesis. As far as one can see, any configuration known to be possible in Oenothera could have arisen in evolution as a result of a series of simple segmental interchanges.

Let us examine some of the consequences of segmental interchange as applied to *O. Lamarckiana*. This species is a compound of the two complexes *gaudens* and *velans*. The known chromosomal configurations are satisfied if the two complexes be assigned the following formulæ:

Velans complex = 1·2, 3·4, 5·8, 7·6, 9·10, 11·12, 13·14
Gaudens complex = 1·2, 3·14, 5·6, 7·4, 12·10, 11·8, 13·9

The two complexes combined form one pair, plus a circle of 12, which may be represented as follows with the *velans* chromosomes in bold face type:

1·2 **3·4** — 4·7 — **7·6**—6·5 — **5·8** — 8·11
| | | | |
1·2 3·14—**14·13**—13·9—**9·10**—10·12—**12·11**

The paired chromosomes are alike, at least as to their ends. All the 12 in the circle must be different. In *O. Lamarckiana*, therefore, there are 13 different kinds of chromosomes, at least with respect to their end arrangements. We should expect, accordingly, 13 different primary $(2n + 1)$ types, provided the addition of some of these chromosomes as extras did not render the plant non-viable. They should all be distinguishable if the interchanged segments contained genes capable of causing morphological differences when present in excess. A triploid theoretically should throw 13 rather than 7 primaries. Since, however, a triploid might consist of 2 *velans* complexes plus 1 *gaudens* or 1 *velans* complex plus 2 *gaudens*, the number of primaries with *velans* and with *gaudens* chromosomes as extras would probably depend upon which of the two kinds of triploids was the parent.

Any $(2n + 1)$ type, with a chromosome from the circle as an extra, should show resemblance in certain particulars to two other $(2n + 1)$ types if all the 12 chromosomes from the circle can exist as extras. Except for the primary with the 1·2 chromosome as an extra, it should not be possible to classify the 15 chromosome types into unrelated groups. DeVries' listing of 7 groups of mutants in this species[11] may have been due to certain parts of a given chromosome, when present in excess, masking the external effects of other chromosomal material and to the non-appearance of some of the primaries.

The addition of an extra 1·2 chromosome should produce a trivalent but should not affect the circle. The addition of one of the chromosomes from the circle as an extra should break the circle or alter it by transforming one

of its members into a bivalent. Secondary ($2n + 1$) types, such as those found in Datura in which the extra has like ends, have not been identified with certainty in Oenothera. A 1·1 chromosome as an extra should form a Secondary, with the pair transformed into a circle of 3. An extra 3·3 chromosome, however, should form a secondary with the circle enlarged to 13 members. Tertiary ($2n + 1$) types, in which the extra has come from interchange between non-homologous chromosomes, have appeared in pure lines of Datura. Such types may be expected in Oenothera but have not yet been identified. They would obviously alter the configurations.

Judging from the configurations in Datura in a hybrid between a standard $4n$ and a "B" $4n$ plant, one might expect a tetraploid Oenothera, which contained 2 *velans* plus 2 *gaudens* complexes, to have 14 pairs, or a quadrivalent and a circle of 24, or a quadrivalent and 2 circles of 12, or a quadrivalent and 12 pairs, or various other configurations. An n number of quadrivalents, however, such as are found in pure line Daturas, would not be expected.

The study of a species with 7 pairs of chromosomes, such as *O. Hookeri*, might help in the analysis of other species. Were a tetraploid to be found in this species, it could be crossed to a diploid to give a triploid. The latter should throw all the viable 7 primary ($2n + 1$) types which could be recognized without chromosomal study and used in analyzing complexes and extra chromosomal types in other species. By use of radiation, the formation of new genes could be induced as well as chromosomal interchanges and translocations which would be additional tools for such analysis. By continued back-crossing of the extra chromosomal types from other species to a standard line of this species, the extra chromosome could be gotten into a soma in which all the other chromosomes belonged to the standard line. It should then be possible, from the morphology of this purified ($2n + 1$) form, from the types which it threw in its offspring, and from the chromosomal configurations induced, to analyze the composition of the extra chromosome in question in terms of a standard. Such a study might help in gaining further insight into the mechanisms involved in the evolution of the Oenotheras.

In the present paper we have compared the chromosomal phenomena in Datura and Oenothera and have shown that segmental interchange is a possible basis of circle formation in both genera. In the immediately following paper, we shall offer a specific line of evidence that segmental interchange has been responsible for the evolution of circles in Oenothera.

[1] Cytological work on Oenothera has been aided by a grant to the second author from the Bache Fund of the NATIONAL ACADEMY OF SCIENCES.

[2] Cleland, R. E., and Oehlkers, Fr., *Amer. Nat.*, **63**, 497–510 (1929).

[3] Blakeslee, A. F., *J. Hered.*, **20**, 177–190 (1929).

[4] Blakeslee, A. F., and Cleland, R. E., "Segmental Interchange, a Possible Basis of Chromosomal Attachments in Oenothera."

[5] Cleland, R. E., *Amer. J. Bot.*, **9**, 391–413 (1922).

[6] Renner, O., *Berichte Deutsch. Bot. Gesell.*, **34**, 858–869 (1917); *Zeitschr. indukt. Abst.-u. Vererb.*, **18**, 121–294 (1917); *Flora*, **11**, 641–667 (1918); *Bibliotheca Genetica*, 9 (1925).

[7] Belling, J., and Blakeslee, A. F., these PROCEEDINGS, **12**, 7–11 (1926).

[8] Blakeslee, A. F., and Cartledge, J. L., *Mem. Hort. Soc. N. Y.*, **3**, 305–312 (1927).

[9] Håkansson, A., *Hereditas*, **11**, 129–181 (1928).

[10] Darlington, C. D., *J. Genet.*, **20**, 345–363 (1929).

[11] DeVries, H., *Genet.*, **8**, 233–238 (1923).

INTERACTION BETWEEN COMPLEXES AS EVIDENCE FOR SEGMENTAL INTERCHANGE IN OENOTHERA[1]

BY RALPH E. CLELAND AND ALBERT F. BLAKESLEE

GOUCHER COLLEGE, BALTIMORE, MD., AND CARNEGIE INSTITUTION OF WASHINGTON, DEPARTMENT OF GENETICS, COLD SPRING HARBOR, N. Y.

Communicated February 13, 1930

In the immediately preceding paper, the formation of circles of 4 or more attached chromosomes during diakinesis has been shown to be explainable on the hypothesis of segmental interchange, and certain relations of this hypothesis to phenomena in Oenothera have been pointed out.

The real test of an hypothesis lies in the success of predictions based upon it. We shall show in the present paper that it is possible successfully to predict on the basis of this hypothesis what the chromosome configuration ought to be in certain complex-combinations. The method is to consider complexes in groups of 3, each complex corresponding to a corner of a triangle, while the chromosome configuration which results when two complexes of the triangle are combined is represented by the side of the triangle connecting the two complexes in question. When the configurations are known for two sides of the triangle, it will be found that in some cases at least the configurations possible to the third side are few in number on the basis of segmental interchange. The configuration representing the third side can therefore be predicted, at least to the extent of reducing the possibilities to a few of the 15 configurations theoretically obtainable in the genus. We have first considered triangles, all three sides of which are known, in order to check the possibilities of this method. Finding that the actual configurations are in each case one of the theoretically predictable possibilities (in some cases the only possibility), we have ventured to make predictions in respect to complex-combinations whose configurations are as yet unknown.

For example, take the triangle *excellens–*^{h}*Hookeri–flavens*.[2]

Excellens.^{h}*Hookeri* has a circle of 4, and 5 pairs[3] and ^{h}*Hookeri.flavens* has also a circle of 4, and 5 pairs.[4] We shall consider the chromosome configuration of *excellens.flavens* for the moment as unknown. On the basis of segmental interchange, it is theoretically possible for *excellens.flavens* to have one of 4 configurations only. These are listed in paragraphs (*a*) to (*d*).

(*a*) *Seven pairs.*—*Excellens* and *flavens* both differ from ^{h}*Hookeri* by a single interchange. They might, therefore, have identical arrangement of ends, in which case, however, they would give, when combined with each other, 7 pairs. This theoretical possibility is not actually possible, however, for if *excellens* and *flavens* have the same end arrangement they must give the same configurations when combined with a given third complex. This they do not do when combined with *gaudens* and *velans*.

(*b*) *Circle of 4, 5 pairs.*—The interchanges giving rise to both *excellens* and *flavens* might have involved the same chromosomes, but different ends have been exchanged in the two cases. *Excellens* and *flavens* would then yield with each other a circle of 4, and 5 pairs.

Let ^{h}*Hookeri* = 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14
Let *flavens* = 1·4, 3·2, 5·6, 7·8, 9·10, 11·12, 13·14
Let *excellens* = 1·3, 2·4, 5·6, 7·8, 9·10, 11·12, 13·14

This possibility may be tested by seeing whether the complexes of the triangle as arranged will give the proper configurations when combined with still other complexes. Using *velans* in this case as a tester, it is found that when *excellens* and *flavens* are so arranged as to give with each other a circle of 4, they cannot both give the proper configuration with *velans*. *Velans* gives with *excellens* a circle of 6 and 4 pairs;[3] and with *flavens* 2 circles of 4 and 3 pairs.[5] But no complex that will give 2 circles of 4 with *flavens* can give a circle of 6 with *excellens* when *flavens* and *excellens* bear the relation they do in the above formulæ. Possibility (*b*) must therefore be eliminated.

(*c*) *Circle of 6, 4 pairs.*—One of the 2 chromosomes involved in the interchange resulting in the *flavens* end arrangement might also have been involved in the formation of the *excellens* arrangement. *Excellens.flavens* should then have a circle of 6 and 4 pairs.

Let ^{h}*Hookeri* = 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14
Let *flavens* = 1·4, 3·2, 5·6, 7·8, 9·10, 11·12, 13·14
Let *excellens* = 1·2, 3·6, 5·4, 7·8, 9·10, 11·12, 13·14

With the data at hand, we have not been able as yet to eliminate this possibility. Thus, using *velans*, *punctulans*[2] and *gaudens* as testers (these being the only 3 complexes whose configurations are known with all 3 of the triangle as well as with each other) it is possible to construct formulæ for all 6 complexes such that each will give the correct chromosome configuration with each of the others, as for instance:

Let *velans* = 1·2, 3·4, 5·8, 7·6, 9·10, 11·12, 13·14
Let *punctulans* = 1·4, 3·8, 5·6, 7·10, 9·12, 11·2, 13·14
Let *gaudens* = 1·2, 4·6, 5·11, 7·8, 9·14, 10·12, 13·3

(*d*) *Two circles of 4, 3 pairs.*—*Excellens* and *flavens* might differ from *h Hookeri* by an interchange involving entirely different pairs of chromosomes. They would then give when combined with each other 2 circles of 4 and 3 pairs.

Let *h Hookeri* = 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14
Let *flavens* = 1·4, 3·2, 5.6, 7·8, 9·10, 11·12, 13·14
Let *excellens* = 1·2, 3·4, 5·8, 7·6, 9·10, 11·12, 13·14

With the triangle arranged in this way, the following formulæ for *velans*, *punctulans* and *gaudens* will meet every demand as far as known:

Let *velans* = 1·2, 3·4, 5·10, 7·8, 9·6, 11·12, 13·14
Let *punctulans* = 1·4, 3·10, 5·6, 7·12, 9·8, 11·2, 13·14
Let *gaudens* = 1·2, 3·13, 5·4, 7·14, 9·10, 11·6, 8·12

This is actually a possibility therefore as far as we can tell from available cytological data.

Our conclusion, therefore, in regard to *excellens.flavens* is that this complex-combination must have on the basis of segmental interchange either a circle of 6 and 4 pairs, or else 2 circles of 4 and 3 pairs. It has, as a matter of fact, the latter configuration.[3]

The other sides of this triangle may be similarly tested. Considering *excellens.h Hookeri* as the unknown, the following are the theoretical possibilities for this combination, so far as this triangle is concerned—circle of 4 and 5 pairs; 2 circles of 4 and 3 pairs; 3 circles of 4 and 1 pair; circle of 4, circle of 6 and 2 pairs; circle of 8 and 3 pairs. Using *velans*, *gaudens* and *punctulans* as testers, the only configuration which survives the test is circle of 4 and 5 pairs. That is, when the complexes of the triangle are so arranged that *excellens.h Hookeri* has any of the theoretical possibilities other than a circle of 4, it is impossible to form *velans*, *gaudens* and *punctulans* which will give the correct chromosome configuration with each of the triangle and with each other. On the basis of segmental interchange, therefore, *excellens.h Hookeri* must have a circle of 4 and 5 pairs, nothing else. As a matter of fact, it has this configuration.[3]

The third side of the triangle, *h Hookeri.flavens*, in the same way may have theoretically, as far as this triangle is concerned, a circle of 4, 5 pairs; 2 circles of 4, 3 pairs; 3 circles of 4, 1 pair; circle of 4, circle of 6, 2 pairs; and a circle of 8, 3 pairs. Two possibilities remain after *velans*, *gaudens* and *punctulans* have been used as testers—namely, circle of 4, 5 pairs; or circle of 4, circle of 6, 2 pairs. Of these, the former is the actual configuration.[4]

In like manner, other complex-combinations, whose configurations are already known, have been tested and in each case the actual configuration

is one of the possibilities under segmental interchange. Thus, *velans. stringens* must have 3 circles of 4, 1 pair; or circle of 4, circle of 6, 2 pairs. It has the latter.[5] *h Hookeri.velans* must have a circle of 4, 5 pairs; or a circle of 4, circle of 6, 2 pairs. It has the former.[3]

Such success in making predictions upon the basis of segmental interchange has led us to make predictions in regard to some complex-combinations whose configurations are as yet unknown. Thus, we venture to predict that *h Hookeri.stringens* will be found to have 2 circles of 4 and 3 pairs; that *excellens.stringens* will be found to have 3 circles of 4 and 1 pair, or a circle of 4, circle of 6 and 2 pairs; and that *acuens.flavens* will be found to have 2 circles of 4 and 3 pairs, or a circle of 8 and 3 pairs.[6]

It will be noted that in only 2 of the above 8 examples, has it been possible to reduce the possibilities to a single one. In 6 cases we have had to content ourselves with predicting that the configuration will be one of two possibilities. This is as far as we can go, using only our present knowledge of chromosome configurations as they have been determined in various complex-combinations, and applying these data according to the theory of segmental interchange. There exists, however, a further check which in several of the above cases eliminates one of the two possibilities and makes it possible to reduce the theoretically predictable configurations to one. We are indebted to Dr. A. H. Sturtevant for the suggestion that a study of linkage relations in Oenothera will in some cases lead to eliminations.[7]

The following facts may be applied as tests to the possibilities which have survived in the above examples: Inasmuch as *R* segregates independently of the rest of the complex and of the lethals in *velans.gaudens* (circle of 12), it probably resides in the single independent pair of chromosomes. But on the basis of segmental interchange, chromosomes must have the same end arrangement in order to pair. *R* also segregates independently in *h Hookeri.rubens* (circle of 10),[8] and since *rubens* and *gaudens* are practically identical, presumably also in *h Hookeri.gaudens* (circle of 10). *h Hookeri* and *velans*, therefore, as well as *rubens* and *gaudens* must have the *R* chromosome in common in respect to end arrangement.

Since *R* and *P* segregate independently of each other and of the rest of the complex in *rigens.velans* or *rigens.h Hookeri*, they lie in different and independent chromosomes in *velans* and *h Hookeri*, and these complexes have identical end arrangement in respect to the *P* chromosome, as well as to that containing *R*. But *R* and *P* are linked in *velans.flavens* (2 circles of 4). They must lie, therefore in a single chromosome group in this combination, that is, in one of the circles of 4. A circle of 4 is obtained on the basis of segmental interchange when two chromosomes which have suffered a simple exchange of ends are brought into association with corresponding chromosomes between which no exchange has occurred. There-

fore the *R* and *P* chromosomes of *flavens* differ from the corresponding *velans* and *hHookeri* chromosomes by a single interchange. In *flavens.stringens* *P* segregates independently of the rest of the complex.[9] It must therefore lie in a free pair of chromosomes. It follows that *flavens* and *stringens* have the *P* chromosome in common in respect to end arrangement.

Fr (pollen fertility factor) segregates independently in *albicans.stringens* (circle of 12) and *flavens.stringens* (circle of 4),[9] so that *albicans*, *flavens* and *stringens* must be alike in respect to end arrangement of the *Fr* chromosome. But *albicans*, while having the *Fr* chromosome in common with both *flavens* and *stringens*, has none in common with *hHookeri* (*albicans.hHookeri* has circle of 14). The only one it can have in common with *flavens* is one in respect to which *flavens* differs from *hHookeri*, that is, in the *R* or *P* chromosome, or in the other *flavens* chromosome in the circle formed with *hHookeri*, if *R* and *P* are in the same *flavens* chromosome. (*flavens.hHookeri* has a circle of 4, 5 pairs). *Fr* resides therefore in one of these. But *Fr* and *P* segregate independently of each other in *flavens.stringens*, so *Fr* and *P* lie in different and independent chromosomes in *flavens* and *stringens*, and *stringens* has not one but both of the chromosomes in common with *flavens* in respect to which *flavens* differs from *hHookeri*. Since it is not possible to tell with certainty which of these two *flavens* chromosomes has *R* and which has *P* (that will depend upon the method of interchange) one cannot be certain which of these chromosomes is common to *albicans* and *stringens*. Furthermore, whether *Fr* lies in the same chromosome as *R* in *flavens*, *albicans* and *stringens* will depend upon the loci of *R* and *P*. (*a*) If both or neither of these two genes are located in the segments which have interchanged, then *R* and *P* will lie in different chromosomes in *flavens*, *Fr* will lie in the *R* chromosome, and it will be the *R* chromosome which *flavens* and *albicans* have in common. (*b*) If, however, one only of these two genes is in the segments which have interchanged, then both *P* and *R* will lie in the same chromosome in *flavens*. The chromosome common to *flavens*, *albicans* and *stringens* then would contain neither *R* nor *P*, and *Fr* would lie in the chromosome lacking both *R* and *P*. It may be possible to determine genetically which is the correct alternative. Meanwhile, *albicans* may be said to have in common with *flavens* one of the 2 chromosomes in respect to which the latter differs from *hHookeri*, and *stringens* has both of these chromosomes in common with *flavens*.

Using these facts as an added check on the predictions obtained above on the basis of purely cytological data, further eliminations must be made in certain cases, owing to the fact that the complexes as they must be arranged in order to give the predicted configurations do not have the proper chromosomes in common as called for by the genetic data. Thus,

h Hookeri.flavens cannot have a circle of 4, circle of 6 and 2 pairs; but only a circle of 4 and 5 pairs (which it actually has). *Velans.stringens* cannot have 3 circles of 4 and 1 pair, but only a circle of 4, circle of 6 and 2 pairs (which it actually has). *h Hookeri. velans* cannot have circle of 4, circle of 6 and 2 pairs, but only a circle of 4 and 5 pairs (which it actually has).[3]

We may illustrate with *h Hookeri.velans*. If this combination has a circle of 4, circle of 6 and 2 pairs, the formulæ for the complexes of the triangle *h Hookeri–flavens–velans* would be as follows:

Let *h Hookeri* = 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14
Let *flavens* = 1·4, 3·2, 5·6, 7·8, 9·10, 11·12, 13·14
Let *velans* = 1·4, 3·6, 5·2, 7·10, 9·8, 11·12, 13·14

According to genetic data, *velans* must have in common with *h Hookeri* the 2 chromosomes in respect to which *h Hookeri* differs from *flavens* (the *R* and *P* chromosomes). No arrangement of *velans* which will give 2 circles of 4 with *flavens* and a circle of 4, circle of 6 with *h Hookeri*, however, can satisfy these conditions. Consequently, *h Hookeri.velans* cannot have a circle of 4, circle of 6 and 2 pairs. It can, however, have a circle of 4 and 5 pairs.

The end result of this attempt at prediction, therefore, is that in the case of 5 hybrid combinations definite predictions have been made as to the chromosome configuration (*excellens.h Hookeri*, *h Hookeri.flavens*, *velans.stringens*, *h Hookeri.velans*, *h Hookeri.stringens*), while less definite predictions have been made in 3 other cases (*excellens.flavens*, *excellens. stringens*, *acuens.flavens*). Of the 5 in which definite predictions have been made, 3 were already known cytologically when the predictions were worked out, and the predictions agreed with the facts, one has since been ascertained to be in line with the prediction (*h Hookeri.velans*) and one still remains to be determined (*h Hookeri.stringens*).

A list of hypothetical formulæ for several complexes which meet all cytological demands on the basis of segmental interchange, (i.e., each gives the correct chromosome configuration with each of the others) and which meet also all genetical demands as far as we have been able to determine them is given herewith.

Let *h Hookeri* = 1·2, 3·4, 5·6, 7·8, 9·10, 11·12, 13·14.

Let 1·2 contain *R*, let 3·4 contain *P*, let 5·6 contain *Sp*, *let*. Then *flavens* must have 1·4, 3·2 or 1·3, 4·2, the rest of the complex being like *h Hookeri*.

Let *flavens* = 1·4, 3·2, 5·6, 7·8, 9·10, 11·12, 13·14.

Then *velans* must have 1·2, 3·4 and differ from *h Hookeri* by one other interchange. Selecting the third and fourth chromosomes arbitrarily (neither *h Hookeri* nor *velans* carries *sp* nor *let*, hence the correctness of this choice is not known).

Let *velans* = 1·2, 3·4, 5·8, 7·6, 9·10, 11·12, 13·14.

Excellens must have 1·2, 3·4 in order to give a circle of 4 with ^{h}*Hookeri*, and 2 circles of 4 with *flavens*. The interchange must involve 5·6 or 7·8 (not both) in order to give a circle of 6 with *velans*. Choosing arbitrarily 7·8 and 9·10 for the interchange,

Let *excellens* = 1·2, 3·4, 5·6, 7·10, 9·8, 11·12, 13·14.

Stringens has 1·4, 3·2 like *flavens*. To give 2 circles of 4 with ^{h}*Hookeri*, a circle of 4 with *flavens*, and a circle of 4, circle of 6 with *velans*, an exchange must have occurred between either 5·6 or 7·8, and one of the last 3 chromosomes. Choosing 5·6 and 9·10 arbitrarily,

Let *stringens* = 1·4, 3·2, 5·10, 7·8, 9·6, 11·12, 13·14.

Rubens and *gaudens* have 1·2 like *velans* and 5·6 like ^{h}*Hookeri* (*Sp*, *sp* segregate independently in ^{h}*Hookeri.rubens*). As nothing is known respecting the position of the genes in the other chromosomes, they may be arranged tentatively in any way to give the correct configurations with the other complexes.

Let *rubens* and *gaudens* = 1·2, 3·14, 5·6, 7·4, 12·10, 11·8, 13·9.

Albicans must have 1·4 or 3·2 like *flavens* and *stringens*. The other chromosomes may be arranged tentatively to give the proper configurations with other complexes.

Let *albicans* = 1·4, 3·5, 6·8, 7·10, 9·12, 11·14, 13·2.

The neatness with which one is able to predict the chromosome configurations in various complex-combinations on the basis of segmental interchange leads to the conclusion that the phenomenon of segmental interchange is probably at the basis of circle formation in Oenothera. A fuller discussion of segmental interchange as applied to Oenothera is soon to appear elsewhere.

[1] Aided by a grant to the first author from the BACHE FUND of the NATIONAL ACADEMY OF SCIENCES.

[2] *Excellens* and *punctulans* are unpublished names applied tentatively by Renner to the egg and pollen complexes, respectively, of the "*biennis* Chicago" of de Vries (*Chicagoensis* of Renner). We are indebted to Professor Renner for permission to use these names.

[3] Unpublished data.

[4] Cleland, R. E., *Zeitschr. indukt. Abst.- u. Vererb.*, 1928, Suppl. Band **1**, 554–567 (1928).

[5] Cleland, R. E., and Oehlkers, Fr., *Am. Nat.*, **63**, 497–510 (1929).

[6] Gerhard assigns to *acuens. flavens* a circle of 4 and 5 pairs (*Zeitschr. Naturwiss.*, **64**, 283–338 (1929)), but inasmuch as his *grandiflora* seems to have a different chromosome configuration from that described by Cleland and Oehlkers, the material used in the two cases is probably not comparable.

[7] In a letter received from Dr. Sturtevant after the above predictions had been made, we learn that he and Dr. S. H. Emerson have together arrived independently at some of the same predictions that we have made.

[8] Renner, O., *Bibliotheca Genetica*, **9** (1925).

[9] Oehlkers, Fr., *Jahrb. wiss. Bot.*, **65**, 403–446 (1926).

POLLEN-TUBE GROWTH OF THE PRIMARY MUTANT OF DATURA, ROLLED, AND ITS TWO SECONDARIES[1]

BY J. T. BUCHHOLZ AND A. F. BLAKESLEE

UNIVERSITY OF ILLINOIS, URBANA, ILL., AND CARNEGIE INSTITUTION OF WASHINGTON, STATION FOR EXPERIMENTAL EVOLUTION, COLD SPRING HARBOR, NEW YORK.

Communicated January 24, 1930

Rolled (Rl) is a primary ($2n + 1$) type in which the largest chromosome (L) is extra. Its two secondary types have been discovered, viz., Sugarloaf (Sg) in which the extra consists of half of the Rl chromosome doubled, represented as the unshaded half in our diagram (Fig. 1); and Polycarpic (Py) which has the other, or shaded half, doubled to form the extra chromosome. The differences in the growth habit and morphological characters of these types as they grow in the garden are very marked. They have been illustrated and described in more or less detail elsewhere[2, 3], and are such that vegetative characters of the two secondaries are complementary to each other in respect to those of their primary Rolled, which are intermediate. In this paper we shall show that the pollen-tube growth conditions in these three types are likewise very characteristically different.

The pollen-tube behavior in normal diploid plants has been described in other publications,[5] more particularly in a paper on balanced chromosomal types.[6] All of the present studies were made from dissected preparations of normal diploid styles which were pollinated with the pollen to be tested, kept at a constant temperature around 18°C. and killed after 12-, 15- or 18-hour intervals. The ends of the pollen tubes, which were growing from left to right, were counted and plotted in the same manner as previously described with reference to the stigma at the left, except that we are using 2-mm. intervals in our diagrams. The number of swollen and burst pollen tubes were plotted below the datum line, the number of normal-appearing pollen tubes above this line, and the number of ungerminated pollen grains remaining on the stigma were represented by the heavy vertical bar at the left. The methods of dissection and other details of technique have been described elsewhere.[5]

Figure 2 represents a distribution of the ends of the pollen tubes of Rolled growing on a $2n$ pistil for a period of 12 hours. As shown in our diagram (Fig. 1) there are two classes of pollen grains, having chromosomal formulas n and (n + Rl) in the proportion 1:1. Under favorable conditions about 96 per cent of these pollen grains may germinate. The first class grows normally and may reach the ovary to accomplish fertilization, while the class (n + Rl) which would produce Rolled plants if they entered an ovule containing an n egg, are not successful. These are represented by the abnormal-appearing pollen tubes plotted below the datum line. These

gametophytes grow slowly, become swollen and burst in the style, usually in the region only a few millimeters below the stigma. The two classes of pollen tubes are easily distinguishable, and the graphic diagram of figure 2 shows that they occur approximately in equal proportions. In similar tests which have been made for 18 hours or longer, there is no great change in the position of the (n + Rl) pollen tubes, while the n group moves forward at a fairly uniform rate. The tubes which have burst are certainly not capable of accomplishing fertilization while those which are only swollen may continue to grow slowly, but ultimately

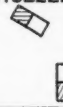
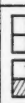
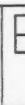
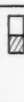
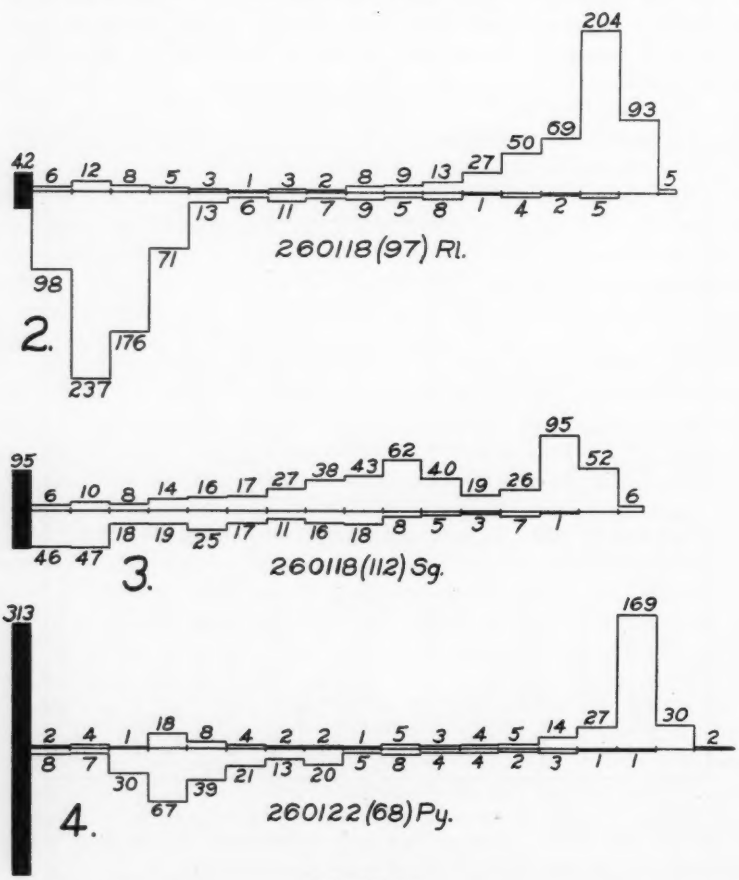
	PRIMARY 2n+1 TYPE		SECONDARY 2n+1 TYPES							
	ROLLED (Rl)		SUGARLOAF (Sg)				POLYCARPIC (Py)			
Trisome in Sporophyte										
Disjunction of Trisome at Reduction										
Formula of ♂ gametophyte	n + Rl	n	n+Sg	n	n+Rl	n-Rl +Sg	n+Py	n	n+Rl	n-Rl +Py
Theor. Product: n egg x this ♂ gametophyte	Rl	Nor.	Sg	Nor.	Rl	?	Py	Nor.	Rl	?
Behavior of Pollen Tubes	Pollen tubes swell and burst near stigma	Pollen tubes grow normally	Pollen tubes grow normally at slower rate	Pollen tubes grow normally	Pollen tubes burst	Aborted pollen grains	Pollen fails to germinate	Pollen tubes grow normally	Pollen tubes burst	Aborted pollen grains

FIGURE 1.

meet the same fate. Tests of this kind explain why the Rl type is not transmitted through the pollen in male back-crosses.

Sugarloaf represents a secondary of Rolled, a ($2n$ + 1) type in which the extra chromosome is made up of the doubled half of the Rolled chromosome. We have represented this Sg half as the unshaded portion in figure 1. Since this trisome has two similar Rl chromosomes and one Sg chromosome which differs from the other two, the microspores produced belong to four classes as shown in the diagram. If distribution of these chromosomes to the poles at reduction occurred at random, these four classes would be in the proportions 2:2:1:1, with the chromosome

formulas n, $(n + Sg)$, $(n + Rl)$ and $(n - Rl + Sg)$. The breeding records as well as the percentage of pollen abortion give evidence that these four classes of gametophytes are produced, but indicate that the classes $(n +$



FIGURES 2–4.

Diagrams of pollen tube distributions in tests of the pollen of Rl, Sg and Py. Only the ends of the pollen tubes were counted and plotted; the spaces on the datum line are 2-mm. intervals.

Rl) and $(n - Rl + Sg)$ are not as abundant as would be expected from random assortment.

The class of microspores $(n - Rl + Sg)$ are entirely deficient for the shaded half of the Rolled chromosome, and these abort without maturing

fully as pollen grains. All plants or individual branches deficient for the whole or a part of a chromosome, thus far recognized in our cultures, have had 50% or more aborted pollen, a fact which shows that pollen grains with chromosomal deficiencies abort. The proportion of aborted pollen cannot be accurately determined from the preparations of our pollen-tube growth tests, but pollen counts made from samples of Sg taken over several seasons have shown that the proportion of aborted grains varies somewhat from time to time, usually averaging between 12.5 and 16 per cent,[4] while $(2n + 1)$ types similar to Rl give from 3 to 6 per cent aborted pollen.

Since the aborted pollen is left out of consideration, we have in figure 3, three classes of matured pollen grains or the pollen tubes produced by them. The normal or n class may be recognized as the group forming the mode of the longest tubes at the right. The Sugarloaf-producing pollen tubes, the $(n + \text{Sg})$ class, form a second mode of normal-appearing tubes, which grow at about $^2/_3$ the rate of the n pollen tubes and constitute a group about equal in number. The pollen tubes whose formula is $(n + \text{Rl})$ constitute the third group plotted below the datum line in the position in which we find the $(n + \text{Rl})$ tubes in figure 2.

The proportion of ungerminated pollen grains remaining on the stigma, 10–11 per cent, is also greater in tests of the pollen of Sugarloaf, than in those of Rolled. Whether this excess represents certain of the $(n - \text{Rl} + \text{Sg})$ deficiencies which matured as pollen grains, or ungerminated grains belonging to the other three classes, is not determinable, but in any event the excess of ungerminated pollen is usually too small to represent the whole of any one of the four classes of microspores. About 4–8 per cent of the pollen of a $2n$ plant fail to germinate. In figure 2 the percentage of ungerminated pollen is 3.5; in figure 3 it is 11.

The tests of the pollen from Sugarloaf showed us first in 1923 that there is a possibility of pollen transmission of this chromosome.[7] It may be seen from these growth tests that the $(n + \text{Sg})$ pollen tubes grow successfully, though at a much slower rate than the normal n tubes, and if the amount of pollen used in making pollinations is not excessive, it should be possible to obtain Sg plants by pollen transmission.

We have tested this prediction genetically by making pollinations with small and medium amounts of pollen from Sg plants onto normal $2n$ plants. The total number of pollen grains applied to the stigma in a pollination of this kind should be limited to 450–600 grains.[7] From a pollination of this kind made in 1927, a pedigree of 59 plants (270965) included 12 Sg plants, while pollinations made in the same way but with styles cut off after the first group of pollen tubes had entered the ovary, contained only normal plants. In routine crosses too much pollen is usually applied to the stigma to transmit more than an occasional Sg

plant through the pollen. Restricted pollination offers a technique by which we may hope to transmit some of the other $(2n + 1)$ types through the pollen.

This genetical test, in which Sg was transmitted through the pollen, serves further to identify the slow-growing group of pollen tubes found in the second mode as the $(n + Sg)$ tubes.

Additional evidence that the slower growing group of pollen tubes carries the double half Sg chromosome is afforded by the behavior of our translocated Sugarloaf type which was obtained by radium treatment. In the type homozygous for the translocation, each pollen grain contains a full complement of chromosomes plus a single attached Sg half. Pollen tubes from such a type form only a single mode, but this mode is intermediate in position between that formed by normal pollen tubes and that formed by the slower growing types represented in figure 3 when grown under the same conditions. As one would expect, the presence of one extra Sg half chromosome slows down the rate of pollen-tube growth less than the presence of two extra Sg halves.

Polycarpic represents a $(2n + 1)$ type with the other half (shaded in Fig. 1) doubled in the extra chromosome. Its trisome is diagrammatically represented in figure 1. It also has the possibility of producing four classes of microscopores: n, $(n + Py)$, $(n + Rl)$, and $(n - Rl + Py)$. There is a high proportion of aborted pollen in this case, usually around 21 per cent, probably representing in part the microspores which are deficient in not having the Sg half among their 12 chromosomes. The other three classes mature as pollen grains and may be accounted for in figure 4. Here the n tubes which constitute the forward mode are the only ones which appear to grow normally. The abnormal pollen tubes include the $(n + Rl)$ tubes which burst in the upper region of the style. The large proportion of ungerminated pollen on the stigma seems to represent the $(n + Py)$ group.

The proportion of burst pollen tubes is somewhat excessive in the pollen of Polycarpic. This may be due to some $(n + Py)$ pollen which actually germinates and then bursts within the style. No genetic tests of pollen transmission of Polycarpic have been successful; our tests show that this would not be expected.

We have selected the primary type Rl, and its secondaries, Sg and Py, for this comparative description because these types involve the largest chromosome of Datura Stramonium in the trisome, and because the secondaries, representing doubled opposite ends of the primary, are known to be extreme types easily recognized in our cultures. These types also represent extreme types with respect to their pollen-tube behavior. The study of these three types has the further advantage of illustrating many of the conditions of pollen-tube growth ordinarily met with among the

other ($2n + 1$) types. Camera lucida drawings of the chromosomes of these three types are reproduced elsewhere.[3]

About 80 tests of each of Rl and Sg have been made at various times in the past seven seasons from field-grown plants. The ones selected for this description are typical of the others, when the pollen is used at a favorable time in the growing season. Very early in the growing season before seed capsules of Rl plants have begun to set and at periods of extreme drought, a condition is sometimes found in which the (n + Rl) pollen grains fail to germinate. Sg also varies slightly in the proportion of its pollen which germinates and sometimes considerably in the distance between the two forward modes at different times in the growing season.

The pollen of Polycarpic has not been available for a large series of tests. We have not made more than a dozen of these in all, and usually from potted greenhouse plants. Polycarpic sets seed capsules in the greenhouse, but almost never when growing in the open. Fig. 4 is fairly representative of all of the tests of Py.

The Rl type cannot be transmitted through pollen. Wherever a male gametophyte has the chromosome formula (n + Rl), the resulting pollen tube appears abnormally swollen or bursts somewhere in the region close to the stigma, whether this Rl-producing pollen grain was derived from a Rl parent, or from the pollen of its secondaries, Sg or Py.

The two halves of the Rl chromosome are not similar in their effect on pollen-tube growth. This is shown by the comparison of the test of the two secondaries, Sg and Py. While the doubled Sg half of the Rl chromosome does not seriously hinder pollen-tube growth, except for the fact that the tubes grow much more slowly, so that (n + Sg) tubes may give rise to Sg plants, the doubled Py half of this Rl chromosome prevents germination of the pollen under the same conditions. Rolled, the primary, stands between the extremes represented by its secondaries. While the (n + Rl) pollen grains usually germinate satisfactorily, the pollen tubes coming from this pollen swell and burst after a short period of growth, a condition which is clearly intermediate between the behavior of (n + Sg) pollen and (n + Py) pollen. The pollen-tube behavior of the Rl group offers an interesting addition to the list of features wherein a primary stands between its two secondaries.[2,3]

[1] This coöperative investigation was made possible by grants from the Joseph Henry Fund of the NATIONAL ACADEMY OF SCIENCES.

[2] Blakeslee, A. F., *Annals., N. Y. Acad. Sci.*, **30**, 1–29 (1927).

[3] Blakeslee, A. F., *J. Hered.*, **20**, 177–190 (1929).

[4] Blakeslee, A. F., and Cartledge, J. L., *Proc. Nat. Acad. Sci.*, **12**, 315–323 (1926).

[5] Buchholz, J. T., and Blakeslee, A. F., *Mem. Hort. Soc. New York*, **3**, 245–260 (1927).

[6] Buchholz, J. T., and Blakeslee, A. F., *Genetics*, **14**, 538–568 (1929).

[7] Davenport, C. B., *Ann. Reports, Director, Dept. of Genetics, Yearbook Car. Inst. Wash.*, **22**, 92 (1924), and **25**, 43 (1926).

OPERATIONAL CALCULUS IN QUANTUM MECHANICS. SOME CRITICAL COMMENTS AND THE SOLUTION OF SPECIAL PROBLEMS

BY R. B. LINDSAY AND R. J. SEEGER

DEPARTMENT OF PHYSICS, YALE UNIVERSITY

Communicated January 17, 1930

The use of operational methods in quantum mechanics goes back to Schrödinger[1] and Eckart[2] who showed that one can always proceed to the wave equation appropriate to a given problem from the corresponding Hamilton-Jacobi equation (in rectangular coördinates) by replacing $\frac{\partial S}{\partial x_i}$ by the operator $\frac{\partial}{\partial x_i}$ considered as operating on the field scalar ψ. And indeed this operator method was used by Schrödinger to establish the relation between his theory and that of Heisenberg. Born and Wiener[3] also employed operational methods in applying quantum mechanics to aperiodic motions. But perhaps the most elaborate use of operational analysis in quantum mechanics is to be found in the work of Hilbert, v. Neumann and Nordheim.[4] It is the purpose of the present paper to comment critically on some points of this theory and also to apply it to the solution of some special problems. It is often in this way that a better understanding of a general point of view is made possible and difficulties that lie concealed in the general presentation are brought to light.

The theory here considered is based on the fundamental assumption that in place of strict functional relations among physical quantities we must remain contented with probability relations. The relations in which we are mainly interested are those among functions of q and p, the conjugate coördinate and momentum, respectively, of a dynamical system (we confine ourselves to one degree of freedom). Thus consider, for example, an electron in an atom. Insofar as we can consider the electron a separate entity from the atom as a whole, it will exist in states having discrete energy values, W_n. What we now assume is that there is no longer a definite orbit associated with each W_n, i.e., a strict functional relation between the distance from the nucleus and the time, but merely a certain *probability* that the state W_n shall lie in a given interval with respect to the nucleus. Physically speaking, for the case of one degree of freedom the idea is equivalent to the distribution of the electronic charge throughout the space occupied by the orbit. As a matter of fact, this latter scheme has been used with some success in the attempt to make calculations on the old Bohr theory of the distribution of electrons in poly-electronic atoms.[5]

The theory we are considering has the advantage that it is cast in postulational form, i.e., the physical requirements that the above-mentioned probabilities must satisfy are set down in a set of postulates. Then a mathematical structure is sought out which will fulfill these requirements, if possible uniquely. As a matter of fact, this latter object is not attained (vide infra). The first physical postulate is the most important for our purposes. It states that corresponding to every pair of dynamical variables $Q_1(p, q)$, $Q_2(p, q)$ there exists a function $\psi(x, y)$ such that $\psi(x, y)\cdot \bar{\psi}(x, y)dx = w(x, y)dx$ is the relative probability that if Q_2 has the value y, Q_1 will lie in the interval $(x, x + dx)$. The function ψ is called the *probability amplitude*. The reader will refer to the original article for the other postulates. Suffice it to say that an operational calculus appears to be the best mathematical structure for the satisfaction of the postulates.

An essential feature of this calculus is the use of *integral* operators. Thus given a function of y, as $f(y)$, this may be transformed into a function of x given by the following integral

$$\int \psi(x, y) f(y)dy = T_\psi f(y), \tag{1}$$

where the integration is extended over the whole range of values of y. In this expression T_ψ appears as a linear operator defined by the identity. The function $\psi(x, y)$ is known as the *kernel* of the operator. In any operator calculus it is essential to define an *idem* operator and an *inverse* operator. The former will be found in the use of Dirac's δ function.[6] Thus we have

$$\mathbf{1}f(y) = \int \delta(x - y) f(y)dy = f(x). \tag{2}$$

The inverse operator to T_ψ is then found in the usual way

$$T_\psi T_\psi^{-1} = \mathbf{1}. \tag{3}$$

Let us note the two following simple functional relations dependent on the behavior of Dirac's δ function, which is zero everywhere for $x \neq y$ and equal to unity for $y = x$, viz:

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right)\delta(x - y) = 0 \tag{4}$$

$$(x - y)\,\delta(x - y) = 0. \tag{5}$$

More relations of this type are possible. We shall discuss the latter below. For the moment we follow the theory as originally developed.

There are now introduced the fundamental operators $q = x\cdot$ (i.e., operation of multiplication by x, the position coördinate) and $p = \epsilon \dfrac{\partial}{\partial x}$, where $\epsilon = h/2\pi i$. The former has the kernel $x\delta(x - y)$ and the latter $\epsilon\delta'(x - y)$.

These operators satisfy by inspection the fundamental condition

$$pq - qp = \epsilon \mathbf{1}. \tag{6}$$

and are said to be *canonically conjugate*. Indeed any pair of operators P and Q satisfying the same condition are said to be canonically conjugate. It is evident that $P = T_\psi p T_\psi^{-1}$ and $Q = T_\psi q T_\psi^{-1}$, where T_ψ is any operator, are canonically conjugate.

From eq. (4) and (5) it then follows that the operators P and Q satisfy the equations

$$(Q - y)\,\psi(x, y) = 0 \tag{7}$$

$$\left(P + \epsilon \frac{\partial}{\partial y}\right)\psi(x, y) = 0, \tag{8}$$

which, of course, are not the most general possible equations of this kind, as will be shown.

The physical interpretation of the above operator calculus is as follows. To every physical quantity considered as a function of coördinate and momentum we assign an operator $Q(p,q)$ by replacing in the analytical expression for the quantity the coördinate by q and the momentum by p. If we confine ourselves at first to functions of the form $Q_1(p) + Q_2(p)$ there is no ambiguity in carrying out this process. The problem is then to find the operator T which transforms q canonically into $Q(p,q)$, i.e., such that $TqT^{-1} = Q$. The kernel of this operator, viz., $\psi(x, y)$, is then assumed to represent the relative probability amplitude that if Q has the exact value y, q will lie in the interval $(x, x + dx)$. This replaces the functional relation of Q and x as dynamical quantities. If we desire to relate *any* two arbitrary quantities $Q_1(p, q)$ and $Q_2(p, q)$, the corresponding procedure will be to find the operators T_1 and T_2 which transform Q_1 and Q_2, respectively. It can then be shown that the kernel of the operator $T_1^{-1}T_2$ is the required probability amplitude that for $Q_1 = y$, Q_2 shall lie in the interval $(x, x + dx)$. It is shown that all the postulational requirements are met by these definitions, but that in order that $w(x, y) = \psi(x, y)\,\bar{\psi}(x, y)$ shall always be real and positive, the operators Q and P must be hermitian, i.e., $\psi(x,y) = \bar{\psi}(y,x)$, and the transformation operator must be hermitian orthogonal, i.e., if T^+ is the operator adjoined to T, we must have $T = \pm(T^+)^{-1}$.

The general procedure involved in the solution of any problem in which the connection between Q and q is sought is to solve the two equations (7) and (8) for $\psi(x, y)$. It is here that there arises another question. If the first equation does not suffice to determine $\psi(x, y)$ completely, presumably the second may suffice to complete the determination. But what is the significance of the second equation when $\psi(x, y)$ is completely determined by the first? This is a question which will presumably be

answered if at all in discussing special cases. We shall recur to it later. For the present we note that if the quantity $Q \equiv H(p, q)$, the Hamiltonian and $y \equiv W$, the numerical value of the energy, then the eq. (7) becomes

$$(H - W)\,\psi(x, y) = 0, \tag{9}$$

which is precisely the equation of Schrödinger, and the functions which satisfy it are the eigenfunctions ψ_n corresponding to the eigenwerte W_n, for which and only which (9) has a solution satisfying the fundamental boundary conditions. For all other values of W the only solution of (9) satisfying the conditions is $\psi = 0$, implying on the present point of view that for any value of $W \neq W_n$ the probability that the position coordinate will be found in any x interval vanishes, i.e., states for which $W \neq W_n$ have zero probability and so do not occur.

The Problem of Uniqueness and the Determination of Canonical Conjugates.—Equations (4) and (5) are very special. From the properties of Dirac's δ function it is seen that the more general relations also hold:

$$\left(\frac{\partial^m}{\partial x^m} + \frac{\partial^n}{\partial y^n}\right)\delta(x - y) = 0 \tag{10}$$

$$(x^k - y^k)\,\delta(x - y) = 0, \tag{11}$$

where m and k are any integers and n is an odd integer. Introducing T_ψ (assumed to operate on x alone) we have from (10)

$$T_\psi\left(\frac{\partial x^m}{\partial^m} + \frac{\partial^n}{\partial y^n}\right)T_\psi^{-1}\,T_\psi\,\delta(x - y) = 0. \tag{12}$$

But since $T_\psi\,\delta(x - y) = \psi(x, y)$ there follows

$$\left\{T_\psi \epsilon \frac{\partial^m}{\partial x^m}\,T_\psi^{-1} + \epsilon\frac{\partial^n}{\partial y^n}\right\}\psi(x, y) = 0. \tag{13}$$

From (11) by similar reasoning there follows

$$\left\{T_\psi\,x^k\,T_\psi^{-1} - y^k\right\}\psi(x, y) = 0. \tag{14}$$

If we now set $Q = T_\psi\,x^k\,T_\psi^{-1}$ and $P = T_\psi \epsilon \dfrac{\partial^m}{\partial x^m}\,T_\psi^{-1}$ the fundamental operator equations corresponding to (7) and (8) become

$$(Q - y^k)\,\psi(x, y) = 0 \tag{15}$$

$$\left(P + \epsilon\frac{\partial^n}{\partial y^n}\right)\psi(x, y) = 0. \tag{16}$$

Equations (7) and (8) form a special case of these equations, when $k = n = 1$. The operators P and Q will satisfy a complicated relation, which for $m = k = 1$ reduces to $PQ - QP = \epsilon\mathbf{1}$. A few other special cases are noted here.

(A) $m = 1,\ k = 2$

$$P(PQ - QP) - (PQ - QP)P = 2\epsilon^2\mathbf{1} \tag{17}$$

(B) $m = 2,\ k = 1$

$$(PQ - QP)Q - Q(PQ - QP) = 2\epsilon^2\mathbf{1} \tag{18}$$

(C) $m = 2,\ k = 2$

$$(PQ - QP) - Q(PQ - QP)Q^{-1} = 8\epsilon\mathbf{1}. \tag{19}$$

The argument of simplicity and convenience would dictate the choice $m = k = 1$, though n is thereby still left arbitrary and uniqueness is lost. Again simplicity would urge $n = 1$. While for $m = n = k = 1$ and $Q \equiv H$ and $y \equiv W$, eq. (15) becomes the Schrödinger equation, for $k \neq 1$, we have the more general equation

$$(H - W^k)\,\psi(x, y) = 0. \tag{20}$$

It should be noted, of course, that it is only for $m = k = 1$ that P and Q are canonically conjugate within the meaning of the definition. However, the lack of uniqueness here emphasized doubtless serves to render the theory more flexible for subsequent development.

An important problem that arises in any application of the theory is that of determining canonical conjugates. Given a function Q of p and q we desire to find a function P such that

$$PQ - QP = \epsilon\mathbf{1}. \tag{21}$$

This can in simple cases be done by inspection, but it seems desirable to look for a general method of attack. The form of Q is important. With a view to the most important applications, we shall assume

$$\begin{aligned} Q &= Q_1(p) + Q_2(q) \\ &= \sum_{j=0}^{n} a_j p^j + \sum_{j=0}^{n} b_j q^j. \end{aligned} \tag{22}$$

A simple but unfortunately not generally realizable case may at once be disposed of. Let G_1 and G_2 be the canonical conjugates of Q_1 and Q_2, respectively, and assume that $G_1 = \varphi_1(q)$ and $G_2 = \varphi_2(p)$. It then develops that the required canonical conjugate $P = (G_1 + G_2)/2$.

In general, if P is any function which can be expanded into a convergent series in powers of p and q, it can be proved that[7]

$$\begin{aligned} Pp^n - p^nP &= -\epsilon\left\{\frac{\partial P}{\partial q}p^n + p\frac{\partial P}{\partial q}p^{n-1} + \ldots\ldots + p^n\frac{\partial P}{\partial q}\right\} \\ Pq^n - q^nP &= \epsilon\left\{\frac{\partial P}{\partial p}q^n + q\frac{\partial P}{\partial p}q^{n-1} + \ldots\ldots + q^n\frac{\partial P}{\partial p}\right\} \end{aligned} \tag{23}$$

In eq. (23) the indicated differentiations are to be understood in the formal algebraic sense in accordance with the rule $\frac{dp^n}{dp} = np^{n-1}$, etc. In the differentiation of a product the order of the terms must be carefully preserved and quotients must be written as products with reciprocal signs. A certain ambiguity will always be present, apparently typical of all methods in quantum mechanics. Thus should the product of p and q be written pq or qp? In such cases it is advisable to write $\frac{1}{2}(pq + qp)$ for complete generality. As a matter of fact the *reality* condition (i.e., the condition that the operators must be hermitian) will often dictate such a step.

Applying (23) to (21) and (22) it follows that the problem of finding P for a given Q reduces to the solution of the following partial differential equation (in non-permutable variables) of the first order

$$\begin{aligned} &\sum_{j=1}^{n} b_j\left\{\frac{\partial P}{\partial p}q^{j-1} + q\frac{\partial P}{\partial p}q^{j-2} + \ldots\ldots + q^{j-1}\frac{\partial P}{\partial p}\right\} \\ &- \sum_{j=1}^{n} a_j\left\{\frac{\partial P}{\partial q}p^{j-1} + p\frac{\partial P}{\partial q}p^{j-2} + \ldots\ldots + p^{j-1}\frac{\partial P}{\partial q}\right\} = 1. \end{aligned} \tag{24}$$

So far as the writers are aware, no work has been done on the solution of differential equations in non-permutable variables. However, they desire to suggest the following tentative procedure, which may receive mathematical refinement later. In each case we shall write the corresponding equation in permutable variables and solve it by Lagrange's method. If a solution for the case of non-permutable variables exists at all it must be included in the general solution for the permutable case, and we seek to find it as some combination of special solutions, being guided by the fact that it must be differentiable like a well-ordered function. The method is best illustrated by some actual examples.

First take the case where $Q = p^2$, corresponding to the Hamiltonian for free motion. The fundamental equation (24) then becomes

$$-\left(\frac{\partial P}{\partial q}p + p\frac{\partial P}{\partial q}\right) = 1. \tag{25}$$

The solution for permutable variables is obtained at once as $P = -q/2p +$

$F(p^2)$ and that for the non-permutable case, taking into account the reality condition, may be written, neglecting an arbitrary constant

$$P = \frac{1}{4}(qp^{-1} + p^{-1}q) + F(p^2) \tag{26}$$

where F is an arbitrary function. More generally $P = -(\alpha qp^{-1} + \beta p^{-1}q)/2(\alpha + \beta) + F(p^2)$, where α and β are arbitrary constants, but $\alpha \neq -\beta$.

As a second case take $Q = p^2 + q$, corresponding to motion in a uniform field (freely falling particle). The equation (24) then becomes

$$-\left(\frac{\partial P}{\partial q}\,p + p\,\frac{\partial P}{\partial q}\right) + \frac{\partial P}{\partial p} = 1. \tag{27}$$

The general solution for permutable variables is $P = f(p, p^2 + q)$ and the proper final solution for our case is

$$P = p + F(p^2 + q). \tag{28}$$

A more striking illustration of the method is provided by the case where $Q = p^2 + q^2$, corresponding to the Hamiltonian for simple harmonic motion. The fundamental equation becomes

$$\left(\frac{\partial P}{\partial p}\,q + q\,\frac{\partial P}{\partial p}\right) - \left(\frac{\partial P}{\partial q}\,p + p\,\frac{\partial P}{\partial q}\right) = 1. \tag{29}$$

The general solution for the case of permutable variables is $P = f(p^2 + q^2,$ arc tan $p/q)$. Now the function arc tan pq^{-1} is not differentiable in the usual way when considered as a function of non-permutable variables, that is, d arc tan $pq^{-1} \neq d(pq^{-1})/[1 + (pq^{-1})^2]$. Hence we must find a combination of functions of it and $p^2 + q^2$ which does satisfy this condition. This is found in the following identity

$$\frac{i}{2} \cdot \log \sqrt{\frac{p^2 + q^2}{2}} - \frac{1}{2} \text{ arc tan } p/q$$
$$= i/2 \cdot \log\left(\frac{p + iq}{\sqrt{2}}\right) = P \tag{30}$$

and this is indeed the required canonical conjugate (with, of course, an additive function of $(p^2 + q^2)$), as may be checked by direct substitution. Incidentally $-i/2 \cdot \log\left(\frac{q + ip}{\sqrt{2}}\right)$ will also serve. The general solution proves to be

$$P = \frac{i}{2(\alpha + \beta)} \cdot \left\{\alpha \log\left(\frac{p + iq}{\sqrt{2}}\right) - \beta \log\left(\frac{q + ip}{\sqrt{2}}\right)\right\} + F(p^2 + q^2),$$

where $\alpha \neq -\beta$.

It is felt to be probable that this method may be applicable to other cases though doubtless analytical difficulties will arise.

Special Problems.—We first consider the application to the case of a free particle. The function Q is here the Hamiltonian $H = p^2/2m$, where m is the mass (non-relativity case). From (26) we have for the canonical conjugate $P = -\frac{m}{2}(qp^{-1} + p^{-1}q)$, wherein for the sake of simplicity we omit the arbitrary function of $p^2/2m$. The form chosen can be shown to be hermitian, whereas qp^{-1} and $p^{-1}q$ individually are not. The fundamental equations (7) and (8) with $Q \equiv H$ now become

$$\frac{\epsilon^2}{2m} \cdot \frac{\partial^2\psi}{\partial x^2} - W\psi = 0 \tag{31}$$

$$\epsilon^2 \frac{\partial}{\partial W}\left(\frac{\partial^2\psi}{\partial x^2}\right) - mx\frac{\partial\psi}{\partial x} - \frac{3}{2}m\psi = 0. \tag{32}$$

The general solution of (31) may be written in the form

$$\psi = Af(W)e^{\pm i\sqrt{\frac{-2mW}{\epsilon^2}}x} \tag{33}$$

where $f(W)$ is arbitrary. However, it is determined by eq. (32). For substitution of (33) into (32) yields

$$f(W) = CW^{-1/4} \tag{34}$$

where C is an arbitrary constant. We now apply the normalization condition[3]

$$\int_0^\infty \int_{-\infty}^{+\infty} \psi(x, W)\bar{\psi}(z, W)dx\,dW = 1. \tag{35}$$

The integral proves to be divergent and has to be treated by a special method (indicated in the article of Kennard just referred to). The normalized solution is found to be

$$\psi(x, W) = \sqrt{\frac{\sqrt{2m}}{h}}\,W^{-1/4}e^{\frac{2\pi i}{h}\sqrt{2mW}\,x}. \tag{36}$$

Now if $\psi(x, W)$ is considered a harmonic function of the time with frequency $\nu = W/h$, we write finally

$$\psi(x, W, t) = \sqrt{\frac{\sqrt{2m}}{h}}\,W^{-1/4}e^{2\pi i\left(\frac{W}{h}t + \sqrt{\frac{2m}{W}}\,x\right)} \tag{37}$$

indicating that ψ is propagated as a progressive wave in the negative x direction (the choice of the negative sign in (33) would result in propagation in the positive x direction) with velocity $V = \sqrt{W/2m}$ in agreement

with the De Broglie-Schrödinger notion. Now the probability that for given energy W the position of the particle shall lie in any interval (x_0, x_1) is

$$\int_{x_0}^{x_1} \psi\bar{\psi}dx = (x_1 - x_0)/hV. \tag{38}$$

That is, the smaller the velocity the greater is the probability that the particle will be found in the given interval.

Consider as a second problem the particle in a uniform field. Here we have $H = p^2/2m + kq$, and the corresponding canonical conjugate in its simplest form is $P = p/k$. The fundamental equations (7) and (8) here become

$$\frac{\partial^2\psi}{\partial x^2} + \frac{2m}{\epsilon^2}(kx - W)\psi = 0 \tag{39}$$

$$\frac{\partial\psi}{\partial x} + k\frac{\partial\psi}{\partial W} = 0. \tag{40}$$

Letting $u = \left(\frac{8\pi^2 m}{k^2h^2}\right)^{1/3}(W - kx)$, the first of the above becomes

$$\frac{\partial^2\psi}{\partial u^2} + u\psi = 0 \tag{41}$$

for which the asymptotic solution for large u is[9]

$$\psi = Au^{-1/4}e^{-2/3 \cdot iu^{3/2}}. \tag{42}$$

Inspection shows that the second fundamental equation (here eq. (40)) is satisfied identically by any function of $W - kx$ but by no other function of W save a constant. Hence it appears that the second equation contributes nothing to the problem. This is in marked contrast to the case of the free particle. The normalized solution for the present problem for the case where $kx < W$ is found to be to a first approximation

$$\psi(x, W, t) = \sqrt{\frac{\sqrt{2m}}{h}}(W - kx)^{-1/4}e^{-\frac{4}{3}\frac{\pi i}{kh}\sqrt{2m}W^{3/2}}e^{2\pi i\left(\frac{W}{h}t + \sqrt{\frac{2mW}{h}}x\right)} \tag{43}$$

The corresponding probability that for given W the particle shall be in any interval (x_0, x_1) satisfying the approximation is

$$\int_{x_0}^{x_1}\frac{\sqrt{2m}}{h}(W - kx)^{-1/2}dx = \frac{(x_1 - x_0)}{Vh}\left[1 + \frac{k}{4W}(x_1 + x_0) + \ldots\right] \tag{44}$$

to a first approximation.

The case of the simple harmonic oscillator is interesting but introduces a difficulty. From the discussion above, it is recalled that the canonical

conjugate to the Hamiltonian in this problem is $P = i/2 \log\left(\frac{p+iq}{\sqrt{2}}\right)$. This is a relatively complicated operator and substitution into the eq. (8) yields a result, the interpretation of which has not so far been made clear. Work on this point is in progress.

[1] E. Schrödinger, *Ann. der Phys.*, **79**, 734 (1926).

[2] C. Eckart, *Phys. Rev.*, **28**, 711 (1926).

[3] M. Born and N. Wiener, *Zs. f. Phys.*, **36**, 174 (1926).

[4] *Math. Ann.*, **98**, 1 (1927).

[5] See, for example, R. B. Lindsay, *J. Math. and Phys.*, Mass. Inst. of Tech., **3**, 191 (1924). See also *J. Opt. Soc. Amer.*, **11**, 17 (1925).

[6] P. A. M. Dirac, *Proc. Roy. Soc.*, **113**, 625 (1927).

[7] The proof of (23) for matrices was given by Born and Jordan, *Zs. f. Phys.*, **34**, 858 (1925). The proof for operators follows similar lines, i. e., by induction from simple operations performed on the fundamental relation (6).

[8] See the original paper (referred to in footnote 4). Also E. H. Kennard, *Zs. f. Physik*, **44**, 326 (1927).

[9] See R. H. Fowler and L. Nordheim, *Proc. Roy. Soc.*, **119**, 176 (1928).

PHYSICAL PROBLEMS WITH DISCONTINUOUS INITIAL CONDITIONS

BY H. BATEMAN

DANIEL GUGGENHEIM GRADUATE SCHOOL OF AERONAUTICS, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated February 10, 1930

1. These problems are usually treated by the methods developed by the great French mathematician Jean Baptist Joseph Fourier who died 100 years ago.

In the case when the differential equation is

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2}, \tag{1}$$

and the initial condition

$$v = f(x) \quad \text{when} \quad \tau = 0. \tag{2}$$

Fourier's solution is (*Oeuvres*, t. 1, p. 401 (1888))—*Théorie Amalyrique de la chaleur*, **1822**—

$$v = \frac{1}{\pi} \int_0^\infty e^{-\lambda^2 \tau}\, d\lambda \int_{-\infty}^{\infty} \cos \lambda(x - \xi) f(\xi) d\xi. \tag{3}$$

This expression satisfies the initial condition when it is legitimate to write

$$\lim_{\tau \to 0} v = \frac{1}{\pi} \int_0^\infty d\lambda \int_{-\infty}^{\infty} \cos \lambda(x - \xi) f(\xi) d\xi = f(x). \qquad (4)$$

For a discussion of these steps reference may be made to Carslaw's *Conduction of Heat*, Chapter 3, 1921.

The theory of the Fourier integral shows that at a point where $f(x)$ has a finite discontinuity, the value which the integral (4) may be expected to take is $\bar{f}(x)$, where

$$\bar{f}(x) = \lim_{\epsilon \to 0} \frac{1}{2} [f(x + \epsilon) + f(x - \epsilon)] = \frac{1}{2} [f(x + 0) + f(x - 0)].$$

This may be different from $f(x)$ and so the solution (3) does not cover all possible cases. The same remark is applicable also to Laplace's solution

$$\begin{aligned} v &= (\tau\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-\frac{(x-\xi)^2}{4\tau}} f(\xi) d\xi \\ &= (\pi)^{-1/2} \int_{-\infty}^{\infty} e^{-s^2} f(x + 2s\sqrt{\tau}) ds, \end{aligned} \qquad (6)$$

which is obtained by changing the order of integration in (3). As $\tau \to 0$, $v \to \bar{f}(x)$. It should be mentioned that Fourier (*Oeuvres*, t. 1, p. 421) uses Laplace's formula to find v in the case when $f(x) = 1$ for $+a > x > -a$ and $f(x) = 0$ for $x^2 > a^2$. His expression for v is

$$v = \frac{1}{2} [F(-x - a, \tau) - F(-x + a, \tau)] \qquad (7)$$

where

$$F(\sigma, \tau) = (2/\sqrt{\pi}) \int_{\sigma/2\sqrt{\tau}}^{\infty} e^{-s^2} ds. \qquad (8)$$

It is easily seen that this makes $v = 1/2$ for $x^2 = a^2$. On the other hand, the solution $v = F(|x - x_0|, \tau - \tau_0)$ corresponds to the following initial condition for $\tau = \tau_0$

$$v = 1 \text{ for } x = x_0, \qquad v = 0 \text{ for } x \neq x_0. \qquad (9)$$

Writing

$$S(x - x_0, \tau - \tau_0) = [\pi(\tau - \tau_0)]^{-1/2} e^{-\frac{(x-x_0)^2}{4(\tau-\tau_0)}}, \qquad (10)$$

the complete solution for the initial condition (2) is

$$v = \int_{-\infty}^{\infty} S(x - x_0, \tau) f(x_0) dx_0 + \sum_{n=1}^{m} F(|x - x_n|, \tau) [f(x_n) - \bar{f}(x_n)], \qquad (11)$$

where the summation extends over all the points of discontinuity of $f(x)$.

This result may be interpreted to mean that the solution $S(x - x_0, \tau)$ corresponds to a unit source associated with an element of length dx_0 while

the solution F corresponds to a unit source associated with a point. When $\tau = \tau_0$ the function $S(x - x_0, \tau - \tau_0)$ is zero at all points except $x = x_0$, where it is infinite, it is in fact a type of function which is now called a Dirac function[1] though functions of this type were used by Fourier 100 years ago.[2]

$S(x - x_0, \tau - \tau_0)$ will be called the first fundamental solution of equation (1) and $F(|x - x_0|, \tau - \tau_0)$ the second fundamental solution. Both of these solutions are defined only for $\tau \geqq \tau_0$. Our aim is to find complete systems of fundamental solutions for the most important partial differential equations of mathematical physics. Only a few preliminary results are given in this paper. The fundamental solutions of the first type have already been studied by several writers[3] and the properties of these solutions are included in the properties of the more general functions that are usually called Green's functions.

2. For the simple wave equation

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} \tag{12}$$

the fundamental solution of the second type which corresponds to the initial condition for $t = t_0$

$$\begin{aligned} u &= 1 \qquad x = x_0 \\ u &= 0 \qquad x \neq x_0 \\ \frac{\partial u}{\partial t} &= 0 \qquad \text{everywhere} \end{aligned} \tag{13}$$

is

$$\begin{aligned} u &= 0 \qquad x > x_0 + t - t_0 \\ u &= 1 \qquad x_0 + t - t_0 > x > x_0 - t + t_0 \\ u &= 0 \qquad x < x_0 - t + t_0 \end{aligned} \tag{14}$$

and this solution is valid for $t \geqq t_0$.

If, on the other hand, the differential equation is

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \tag{15}$$

which corresponds to a propagation of waves in one direction only, the fundamental solution of the second type corresponding to the initial condition $u = 1, x = x_0, u = 0, x \neq x_0$ (for $t = t_0$) is

$$u = 1 \quad x = x_0 + t - t_0, \quad u = 0 \text{ elsewhere.}$$

3. We now pass on to a consideration of the equations

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \tag{16}$$

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \tag{17}$$

Considering first the wave-equation, let u be required to satisfy the following conditions (for $t = 0$)

$u = f(x, y, z)$ on the sphere S whose equation is $x^2 + y^2 + z^2 = a^2$, $u = 0$ elsewhere, $\frac{\partial u}{\partial t} = 0$ everywhere.

An appropriate solution has been found elsewhere[4] to be

$$u = \frac{a}{r}\bar{f}, \tag{18}$$

where $r^2 = x^2 + y^2 + z^2$ and $\bar{f}$ is the mean value of f along that circle C on S whose points are all at distance t from (x, y, z). If no such circle exists the value of u is zero.

A solution of equation (16) can often be derived from a solution of (17) by making use of the theorem that if $u(x, y, z, t)$ is a solution of (17) answering the requirements that $u = G(x, y, z)$, $\frac{\partial u}{\partial t} = 0$ when $t = 0$, then

$$v = (2/\sqrt{\pi}) \int_0^\infty u(x, y, z, 2s\sqrt{\tau})e^{-s^2}\, ds \tag{19}$$

is usually a solution of (16) which satisfies the requirement

$$v = G(x, y, z) \text{ when } \tau = 0. \tag{20}$$

This result may be used to find $v(x, y, z, \tau)$ in the case when G is an assigned function of position for points on the sphere $x^2 + y^2 + z^2 = a^2$ and is zero elsewhere. In the particular case when v is required to be unity over the sphere S at time $\tau = 0$ and zero elsewhere at this time, the solution is found by this method to be

$$v = \frac{a}{r}\left[F\{|a - r|, \tau\} - F\{(a + r, \tau\}\right]. \tag{21}$$

The general problem of this type for the equation of conduction arises when v has initially an assigned value $f(x, y, z)$ for each point (x, y, z) of a certain surface S and is initially zero elsewhere. The solution of this problem is evidently unique because if there were two different solutions of the problem their difference would be a non-vanishing solution of (16) with an initial value zero everywhere. Such a solution does not exist.

To determine the value of v it is convenient to consider the definite integral

$$w = \int_0^\infty v(x, y, z, \tau)e^{-\lambda^2\tau}\, d\tau. \tag{22}$$

If the differentiations under the integral sign are permissible this integral satisfies the differential equation

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} - \lambda^2 w = -v(x, y, z, 0). \tag{23}$$

Let us consider the case when $v(x, y, z, 0) = Y_n(\theta, \phi)$ where Y is a spherical harmonic of degree n. Taking the surface S to be the sphere $x^2 + y^2 + z^2 = a^2$ and assuming that $w = R(r).Y_n(\theta, \phi)$, we find that at points not on S, the function $W(r) = rR(r)$ must satisfy the equation

$$\frac{d^2W}{dr^2} = \left[\frac{n(n+1)}{r^2} + \lambda^2\right]W. \tag{24}$$

Let W_0 be a solution suitable for the space outside S, W_1 a solution suitable for the space inside S and let the arbitrary constants in W_0 and W_1 be adjusted so that

$$W_0\frac{dW_1}{dr} - W_1\frac{dW_0}{dr} = 1. \tag{25}$$

We now replace S by a thin shell bounded by the surfaces $r = a - \epsilon \equiv \alpha$ and $r = a + \epsilon \equiv \beta$. Assuming that

$$\begin{aligned} rw &= AW_1(r)Y_n(\theta, \phi) && r < \alpha \\ rw &= [PW_0(r) + QW_1(r) + H(r)]Y_n(\theta, \phi) && \alpha < r < \beta \\ rw &= BW_0(r)Y_n(\theta, \phi) && r > \beta \end{aligned}$$

the continuity of rw and $\frac{\partial}{\partial r}(rw)$ require that

$$\begin{aligned} AW_1(\alpha) &= PW_0(\alpha) + QW_1(\alpha) + H(\alpha) \\ AW_1'(\alpha) &= PW_0'(\alpha) + QW_1'(\alpha) + H'(\alpha) \\ BW_0(\beta) &= PW_0(\beta) + QW_1(\beta) + H(\beta) \\ BW_0'(\beta) &= PW_0'(\beta) + QW_1'(\beta) + H'(\beta). \end{aligned}$$

These equations give

$$\begin{aligned} A &= Q + W_0(\alpha)H'(\alpha) - W_0'(\alpha)H(\alpha) \\ O &= Q + W_0(\beta)H'(\beta) - W_0'(\beta)H(\beta) \\ A &= \int_\alpha^\beta [W_0''(r)H(r) - W_0(r)H''(r)]dr \end{aligned}$$

$$\text{i.e., } A = \int_\alpha^\beta W_0(r)\left[\left\{(\lambda^2 + \frac{n(n+1)}{r^2}\right\}H(r) - H''(r)\right]dr.$$

Similarly

$$B = \int_{\alpha}^{\beta} W_1(r)\left[\left\{\lambda^2 + \frac{n(n+1)}{r^2}\right\} H(r) - H''(r)\right]dr.$$

To ascertain how $H(r)$ must be chosen we consider the case $n = 0$ the solution for which may be derived from (21) by means of equation (22). In this case we find that

$$\begin{aligned} \lambda W_0(r) &= e^{-\lambda r} & \lambda W_1(r) &= \sinh \lambda r \\ \lambda A &= 2ae^{-\lambda a} & \lambda B &= 2a \sinh \lambda a \end{aligned} \tag{26}$$

and the correct result is obtained by writing

$$\lambda^2 H(r) - H''(r) = aX[\lambda(r - a)], \tag{27}$$

where $X(0) = 1$ and $X(s)$ is a function such that for any continuous function $f(s)$

$$\lim_{\epsilon \to 0} \int_{-\lambda\epsilon}^{\lambda\epsilon} f(s)X(s)ds = 2f(0). \tag{28}$$

This function $X(s)$ must, like a Dirac function, become infinite for one or more values of s in the range $(-\lambda\epsilon, \lambda\epsilon)$. In the general case we write

$$\left\{\lambda^2 + \frac{n(n+1)}{r^2}\right\} H(r) - H''(r) = aX[\lambda(r - a)] \tag{29}$$

and the analysis gives simply

$$\lambda A = 2aW_0(a) \qquad \lambda B = 2aW_1(a). \tag{30}$$

The solution of our problem is thus

$$\begin{aligned} \lambda rw &= 2aW_0(a)W_1(r)Y_n(\theta, \phi) & r < a \\ \lambda rw &= 2aW_1(a)W_0(r)Y_n(\theta, \phi) & r > a. \end{aligned} \tag{31}$$

In the case when $n = 0$ the solution is

$$\begin{aligned} \lambda^2 rw &= 2ae^{-\lambda a} \sinh \lambda r & r < a \\ \lambda^2 rw &= 2ae^{-\lambda r} \sinh \lambda a & r > a. \end{aligned} \tag{32}$$

The corresponding solution for the plane $x = x_0$ is

$$\lambda^2 w = e^{-\lambda|x - x_0|}. \tag{33}$$

The solution of the conduction problem for the sphere S may be obtained directly by means of equation (19) and the solution of the wave-equation

$$u = (a/r)P_n(\cos \omega)Y_n(\theta, \phi), \tag{34}$$

where $t^2 = r^2 + a^2 - 2ar \cos \omega$, when this equation gives a real value of ω, u being otherwise zero. The expression for v is thus

$$v = (2/\sqrt{\pi}) \int_{\frac{|r-a|}{2\sqrt{\tau}}}^{\frac{r+a}{2\sqrt{\tau}}} (a/r) P_n \left[\frac{r^2 + a^2 - s^2\tau}{2ar} \right] e^{-s^2}\, ds \cdot Y_n(\theta \cdot \phi)$$

$$= \frac{a^2 Y_n(\theta, \phi)}{\sqrt{(\pi\tau)}} \int_0^{\pi} P_n(\cos\omega) \sin\omega \cdot d\omega \, \frac{1}{\rho}\, e^{-(\rho^2/4\tau)} \qquad (35)$$

where $\rho^2 = r^2 + a^2 - 2ar\cos\omega$. We have also the expressions

$$v = 2a^2(-)^n Y_n(\theta, \phi) \int_0^{\infty} e^{-k^2\tau} J_{n+1/2}(kr_1) J_{-n-1/2}(kr_2) dk/(ar)^{1/2}$$

$$r_1 = r, \quad r_2 = a \text{ if } r < a$$

$$r_1 = a, \quad r_2 = r \text{ if } r > a$$

and this, combined with (31) gives the relation

$$W_0(r_2) W_1(r_1) = (-)^n \int_0^{\infty} \frac{\lambda dk}{\lambda^2 + k^2} J_{n+1/2}(kr_1) J_{-n-1/2}(kr_2)/(r_1 r_2)^{1/2} \qquad r_1 < r_2$$

which is closely related to an integral given by Sonine.[5]

[1] After P. A. M. Dirac, *Proc. Roy. Soc. Lond.*, **113** (1927), 621.

[2] Oeuvres, t. 1, p. 234.

[3] I. Fredholm, *Acta. Math.*, **23** (1900), 1; *Compt. Rend.*, **129** (1899), 32; *Rend. Palermo*, **25** (1908), 346. J. Le Roux, *Compt. Rend.*, **137** (1903), 1230. N. Zeilon, *Arkiv. Mat. Astr. och Fysik.*, **6** (1911), **9** (1914); *Nov. Act. Soc. Sc. Upsaliensis*, **4**, (1919) 5, J. Hadamard, *Lectures on Cauchy's Problem*, New Haven, **1923**, Ch. 3.

[4] *Ann. Math.*, **31** (1930), 158.

[5] N. Sonine, *Math. Ann.*, **16** (1880), 59. See also H. M. Macdonald, *Proc. Lond. Math. Soc.*, **1**, 35 (1902), 428.

ON THE POSSIBLE INFLUENCE OF THE MOSAIC STRUCTURE OF CRYSTALS ON THE DETERMINATION OF AVOGADRO'S NUMBER

By F. Zwicky

Norman Bridge Laboratory, California Institute of Technology, Pasadena

Communicated January 20, 1930

(*A*) *The Experimental Situation.*—The experimental determination of Avogadro's number is very closely related to the determination of the charge of the electron on the one hand and to the absolute determination of the wave-length of x-rays on the other hand. Millikan, as is well known, measured the charge of the electron directly by the oil drop method, long before absolute measurements of the wave-lengths of x-rays were attempted. Values for these wave-lengths, therefore, were obtained from the charge e of the electron in the following indirect way. From e and Faraday's constant F, one derives immediately Avogadro's number,

equal to F/e. The next step is to calculate the lattice constant d of some suitable crystal, whose structure is known, calcite for instance. The spacing d is obtained from the density of the crystal and Avogadro's number. Observing then, the Bragg angle of reflection, Θ for a certain x-ray line one deduces λ from Bragg's relation $2d \sin \Theta = n\lambda$. Corrections for the refractive index of x-rays, etc., have, of course, to be introduced. (See R. T. Birge, *Phys. Review Suppl.*, No. 1, 1929.) For the K_α line of Cu, one obtains in this way $\lambda = 1.5386$ Å.

Recently, methods have been developed to measure the wave-length of x-rays directly by reflecting them on artificially ruled gratings of known spacing. The most accurate value for the above K_α of Cu has probably been obtained by Dr. J. A. Bearden[1] of the University of Chicago. His result is $\lambda = 1.5422$ Å, which is about 0.23% higher than the value of λ mentioned before.

We should mention also some older absolute determinations of the wave-length of x-rays which yielded results less different from those obtained by Bragg's method. Wadlund,[2] in fact, obtained practically identical results for Cu K_α with the two methods. E. Baecklin,[3] on the other hand, found for the K_α line of Al the value $\lambda = 8.333$ Å $\pm$ 0.1%, whereas the value derived from reflections on a gypsum crystal is $\lambda' = 8.3229$ Å, the discrepancy being 0.12%. The values for the charge of the electron and for the Avogadro number as given by Millikan are therefore lower than those derived from the absolute determination of the wave-length of x-rays, the discrepancy being of the order of 0.4% (Baecklin) and 0.7% (Bearden).

(*B*) *On a Possible Explanation on the Basis of the Mosaic Structure of Crystals.*—The question arises, of course, immediately whether the above discrepancies are due to some sort of experimental errors or whether they are of a systematic nature. It seems to the author that a discrepancy between the results obtained by the two methods sketched in the above *necessarily must exist.* In the chain of reasoning which relates the electronic charge to the wave-length of x-rays, there is one doubtful step. Indeed one has to introduce the assumption that the atoms in the crystal are arranged in an absolutely uniform way, being located in the corners of an ideal crystallographic lattice. Now according to the conception of the *mosaic structure*[4] of crystals this assumption is not quite correct. The theory predicts that the mosaic pattern of rock salt, for instance, is represented by a cubic frame work of planes whose density is about 10% larger than the density of the perfect blocks in between them. As the latter cause the interference of the x-rays, their characteristic lattice constant d_0 must be introduced in Bragg's formula. d_0 is obviously larger than the distance d which is obtained in the usual way by assuming that the crystal is of perfectly uniform density throughout. It appears then,

that we really have to expect a too low value of λ, if the indirect determination is used without taking into account the mosaic structure of the crystals.

The theory is not yet advanced enough to allow us to treat accurately the effects due to the mosaic structure. It enables us, however, to estimate them within quite narrow limits. For rock salt, for instance, the planes which define the mosaic pattern contain 10% more atoms per unit area than the regular crystal planes. In order to abbreviate the expressions, I have proposed[4] to denote the mosaic planes with the letter II in contradistinction to the regular p-planes. Suppose now, that we have in a regular cubic crystal, one II plane of density $\rho' = \rho_0(1 + \Delta)$ for n p-planes parallel to it of the density ρ_0, then the average density of the crystal will be

$$\rho = \frac{(n-2)\rho_0 + 3\rho'}{n+1} = \rho_0\left[1 + \frac{3\Delta}{n+1}\right]. \qquad (1)$$

For rock salt, we have $\Delta \cong 0.1$. For the distance D between the II-planes the same argument can be given which I published in another place[5] for the mutual distance of the surface cracks yielding $D \cong 100$ Å. This means that among about 30 parallel p-planes, we will find on the average one II-plane. This gives according to relation (1) a relative correction γ for the density equal to

$$\gamma = \frac{\rho - \rho_0}{\rho_0} \cong 0.01$$

which has to be considered as an upper limit as we have chosen the maximum possible value for Δ. The lattice constant d_0 entering Bragg's relation evidently must be deduced from ρ_0 rather than from the directly observed density ρ. The difference between the two according to the above is

$$d - d_0 \cong 0.01\, d_0/3$$

and the same correction of 0.3% must be applied to the wave-length of x-rays as determined by Bragg's method. Our consideration, therefore, shows that a correction due to the mosaic structure may be of practical importance if the accuracy of the experiments is carried as far as it has been done recently.

(*C*) *On Some Possible Tests of the Theory.*—If the above considerations are correct, one should expect differences for the wave-lengths of x-rays if determined by Bragg's method on different crystals, rock salt and calcite for instance. For the commonly used heteropolar crystals such as NaCl, $CaCO_3$, $CaSO_4$, etc., the effects due to the mosaic pattern will probably be very nearly the same. Assuming that the difference to be expected is of the order of 10% in γ Bragg's method for the determination of the

wave-lengths of x-rays should yield results at variance by about 0.03% in λ if measurements on two different crystals, say calcite and rock salt are compared. In order to check up on our consideration in this manner, it would be necessary to make all measurements accurate to within 10% of the above effects due to the mosaic structure. The necessary operations obviously would be these:

(1) A determination of the density of different samples of two types of crystals, with an accuracy of about one part in 10,000.

(2) A chemical analysis of the same samples with an accuracy of 1 in 10,000.

(3) A determination of the wave-length of a given x-ray line on the two crystals applying Bragg's generalized relation, the accuracy required being 0.003%.

Individual investigations of the above kind have already been carried out with sufficient accuracy in the past. As regards (1), I mention, for instance, the work of O. K. de Foe and A. H. Compton[6] on the density of calcite and of rock salt with an accuracy of 0.1 to 0.2 mg. per cm.^{3} Chemical analyses of crystals with the accuracy required under (2) have also been carried out. It is known, for instance, that Iceland spar of the purest kind contains only 0.03% Fe_2O_3 and no other chemically measurable contaminations. As to the third point, it may be mentioned that the precision measurements carried out in Siegbahn's laboratories are of adequate accuracy. It seems to me, however, that all these investigations are too disconnected for our purpose and do not allow us to draw any definite conclusions. A systematic combined investigation of the three points mentioned under (1), (2) and (3), therefore would be very desirable.

(*D*) *On Other Tests of the Mosaic Structure.*—It may seem more hopeful to determine directly the mosaic structure of the crystals used for the reflection of x-rays and to derive the corrections which have to be applied to the observed wave-lengths afterwards. Investigations with this purpose in mind are now in progress at this Institute, but no definite results have as yet been obtained. The mosaic pattern, although it is of absolute geometrical regularity, at least in good crystals, nevertheless makes a very inefficient grating because of the small difference in reflective power between the II-planes and the p-planes. We have succeeded in showing, for instance, that the mutual distances of the II-planes in perfect metal single crystals are constant within a few per cent, but have not yet been able to observe the diffraction pattern for light reflected from the cross gratings which are formed by the intersection of the mosaic structure with the cleavage faces. The low efficiency of these gratings is probably responsible for the fact that no ghosts of the x-ray lines reflected from crystals have been resolved, though they should theoretically exist.

(*E*) *Effects Due to Internal Cracks.*—In an earlier paper,[5] I have pre-

sented some theoretical reasons for the existence of surface cracks on heteropolar crystals. Both theoretical and experimental evidence can be adduced showing that cracks of the same kind also exist in the interior of these crystals. Different authors[7] have discussed the possibility of clearing up the discrepancy mentioned at the beginning of this paper on these grounds. Most of them came to the conclusion that the existence of internal cracks would make it necessary to introduce corrections which would make the stated discrepancy still worse. It appears to me, however, that no effect is to be expected at all, at least, not if the cracks are of the nature which I have proposed in my paper. The opening of the cracks and the resulting decrease of the average density of the crystal is, indeed, exactly compensated for by the contraction of one or more surface planes bounded by the cracks. The interior structure of the blocks which are formed in this manner, is hardly affected, and it is these blocks, which constitute the major part of the crystal. Their structure, therefore, determines the interference pattern of the reflected x-rays and this will be the same as that given by an ideal crystal except for the intensity and the width of the reflected lines.

[1] J. A. Bearden, *Proc. Nat. Acad. Sci.*, **15**, 528–533 (1929).
[2] A. P. R. Wadlund, *Phys. Review*, **32**, 841 (1928).
[3] E. Baecklin, *Upsala Dissertation*, 1928.
[4] F. Zwicky, *Proc. Nat. Acad. Sci.*, **15**, 816 (1929).
[5] F. Zwicky, *Ibid.*, **15**, 253 (1929).
[6] O. K. de Foe and A. H. Compton, *Phys. Review*, **25**, 618 (1925).
[7] See for instance, A. H. Compton, *J. Frank. Inst.*, **208**, 605 (1929).

THE PROBLEM OF PREPARING OPTICALLY ACTIVE FREE RADICALS

BY EVERETT S. WALLIS

FRICK CHEMICAL LABORATORY, PRINCETON UNIVERSITY

Communicated January 31, 1930

In carrying out certain investigations on the Beckmann rearrangement Jones and Wallis[1] studied the nature of the products formed by the rearrangement of certain optically active acid azides, and optically active hydroxamic acids. *d*-Benzylmethylacetazide, $C_7H_7(CH_3)HCCON_3$, was found to give on rearrangement an optically active isocyanate, $C_7H_7(CH_3)HCNCO$, which could be converted into an optically active amine hydrochloride, and by treatment with aniline into an optically active monosubstituted urea. These experiments together with certain other facts taken from a survey of the properties of optically active compounds

described in the literature led them to suggest that it was very probable that in the rearrangement the optically active radical, if existing as a free radical, did so in such a way as to maintain a configuration required to produce an optically active rearrangement product, that is, it was a positive radical which is of the nature of a carbonium ion, and that in its wandering from the carbon atom to the nitrogen atom it leaves behind the electron which it shared with the carbon atom.

These investigations suggested to the author two problems: 1. A study of the Beckmann rearrangement in the presence of free radicals; 2. The possibility of preparing an optically active free radical. The results obtained from a study of the first problem have been published.[2] I now wish to report on certain investigations which of necessity had to be carried out before any answer could be given to the second problem.

The first step in the preparation of an optically active free radical involves the preparation of certain derivatives of an optically active tertiary carbinol, and since derivatives of tertiary carbinols of the aliphatic series do not form free radicals, we are limited in our choice to tertiary carbinols of the aromatic series containing an asymmetric carbon atom. A survey of the literature shows that no derivatives of tertiary carbinols of this type have hitherto been resolved. Therefore any method of resolution of these compounds should be of interest.

For these investigations I chose phenylbiphenyl-α-naphthyl carbinol, $C_6H_5(C_6H_5C_6H_4)(C_{10}H_7)COH$. This carbinol in the form of its chloride was prepared by the method of Schlenk[3] and had a melting point of 196.5–197°C. The carbinol was also prepared and melted at 162°C.

Several attempts were made to resolve the carbinol itself by combining it directly with brucine, quinine, camphor sulphonic acid, etc., but only negative results were obtained. Experiments were carried out in an attempt to make the carbinol in the form of its chloride combine with silver salts of optically active acids, but this method of approach proved unsatisfactory. Attempts to prepare the half ester of phthalic acid and the carbinol also failed so that method of resolution with an alkaloid could not be used. The chloride did combine with the sodium salt of optically active amyl alcohol as was first reported by Schmidlin[4] but fractional crystallization gave only a very small amount of the optically active amyl ether and this method was abandoned.

Phenylbiphenyl-α-naphthylchlormethane, however, was found to combine with thioglycollic acid according to the following equation

$$C_6H_5(C_6H_5C_6H_4)(C_{10}H_7)CCl + HSCH_2COOH \longrightarrow$$
$$HCl + C_6H_5(C_6H_5C_6H_4)(C_{10}H_7)C\text{-}SCH_2COOH,$$

giving a stable thioglycollic acid derivative which combined with brucine

and was easily resolved, giving the thioglycollic acid derivative of the optically active tertiary carbinol.

Preparation of d,l-Phenylbiphenyl-α-naphthylmethylthioglycollic Acid.—

$$C_6H_5(C_6H_5C_6H_4)(C_{10}H_7)C\text{–}SCH_2COOH,$$

10.90 g. of thioglycollic acid were refluxed with 47.8 g. phenylbiphenyl-α-naphthylchlormethane in 100 cc. dry toluene. The reaction began at a temperature of about 70°C., and hydrogen chloride gas came off as the by-product of the reaction. After the end of the reaction, which usually took about one hour, the solution was allowed to cool, and the white crystalline precipitate which separated was collected at the pump. This was washed with petrol ether, and after two crystallizations from hot toluene gave a white crystalline material which contained one molecule of toluene of crystallization and melted sharply at 116–116.5°C. Yield, 51 g. Analysis for sulphur. Carius method. 0.2006 g. of substance. gave 0.0851 g. $BaSO_4$. Calc., S: 5.79 per cent. Found, 5.82 per cent.

Resolution of Phenylbiphenyl-α-naphthylmethylthioglycollic Acid.—85 g. of the racemic material were dissolved in 200 cc. of dry acetone and added while still warm to 90 g. brucine suspended in 400 cc. of acetone. Solution took place immediately. The mixture was filtered, and allowed to cool overnight without stirring. A large amount of a white material separated. This was collected at the pump, and fractionally crystallized from a hot chloroform petrol ether solution. After four crystallizations no further change in the optical activity or the melting point of the solid could be detected. Forty g. of brucine salt were obtained which melted sharply at 215°C.

1.2124 g. of this material dissolved in 25 cc. of chloroform solution gave a rotation of −1.64 in a 2-dc. tube at 20°C. $[\alpha]_D^{20} = -16.91$.

Preparation of l-Phenylbiphenyl-α-naphthylmethylglycollic Acid.—39 g. of the brucine salt were suspended in ether and decomposed with dilute sulphuric acid. The ether layer was washed free from the sulphate ion and then dried with anhydrous sodium sulphate. This solution was concentrated to about 150 cc. and the acid slowly precipitated by addition to petrol ether. The white material which precipitated was collected at the pump, and recrystallized from warm toluene-petrol ether solution. The dry product melted at 109–110°C.

1.0242 g. dissolved in 25 cc. of dry ether solution gave a rotation of −1.06 at 25°C. in a 2-dc. tube. $[\alpha]_D^{25} = -12.93°$.

1.002 g. dissolved in 25 cc. of carbon tetrachloride solution gave a rotation of −1.09 at 20°C. in a 2-dc. tube. $[\alpha]_D^{20} = -13.63°$. Analysis for sulphur. Carius method.

0.1412 g. of substance gave 0.0722 g. $BaSO_4$. Cal. S: 6.95 per cent. Found, 6.90 per cent.

Conclusions.—A method of resolution of tertiary carbinols of the aromatic series containing an asymmetric carbon atom has been described by means of which an optically active derivative of phenylbiphenyl-α-naphthyl carbinol has been prepared. Investigations are now in progress with a view of converting this compound into an optically active triarylmethyl. The results of these experiments will be reported at a later date and should be of interest not only in connection with the Beckmann rearrangement but also should contribute further valuable information on such important questions as the Walden Inversion and the constitution of the triaryl methyls in general.

[1] Jones and Wallis, *J. Am. Chem. Soc.*, **48**, 169 (1926).
[2] Wallis, *Ibid.*, **51**, 2982 (1929).
[3] Schlenk, *Ann.*, **394**, 196 (1912).
[4] Schmidlin, *Ber.*, **45**, 3188 (1912).

THE IMPLANTATION OF ADDITIONAL HYPOPHYSEAL RUDIMENTS IN URODELE EMBRYOS

BY RAYMOND F. BLOUNT

OSBORN ZOÖLOGICAL LABORATORY, YALE UNIVERSITY

Communicated February 5, 1930

A series of experiments has been performed in which additional hypophyseal rudiments have been implanted in urodele embryos. The object of the transplantation has been to produce a continuous excess of a secretion similar in character and mode of excretion to that of the normal gland. The transplants have been from embryos of the same age, which, by developing concomitantly with the normal gland, bring about a condition comparable to spontaneous hypersecretion which occasionally occurs in animals. This cannot be done by methods formerly used, such as periodic injection, feeding, or by grafting adult glands in late stages. Recently Allen ('29)[1] has grafted pituitary anlagen from the embryo into the larvae of frogs. Since the age of the graft and the age of the host are widely different, neither the method nor the results of this work are directly comparable to those of the present.

Embryos of Amblystoma punctatum were used for the majority of the transplantations, although for some of the work those of A. tigrinum and the axolotl were used. The operations were performed before the invagination of the pituitary rudiment (Harrison's Stages 25–31).[2]

The operations were of two types. The first consisted in removing the ectoderm and underlying neural tube from the region immediately anterior to the stomodeum of one animal and placing it in an opening prepared in

another animal. In most cases the host received two separately placed grafts, in other cases only one, referred to here as triple and double pituitary animals, respectively. The site of the graft was either between the ear and the eye or above and behind the fore limb, the latter position being used exclusively in 1929. In the second type of operation great care was taken to transplant buccal ectoderm only. In other respects the procedure was identical with that followed in the first series.

In each case the operated animal has had its individual controls, one normal and one with extirpated hypophysis; that is, a normal animal and the animals which were used as donors have been raised and compared as regards growth, pigmentation and other differences. The experiments were also controlled by animals into which extra thyroid glands had been grafted. Conditions of environment and food were as nearly identical as possible. Feeding was maximum in the younger animals. Later both the control and the experimental animals were liberally and proportionately fed. The animals were preserved at intervals from the time of operation to some time after metamorphosis. Their controls were killed at the same time. An additional attempt was made to control the work in which the first type of operation was used by checking, in another series of experiments, the function of the grafts. The entire pituitary region was removed from a series of animals and grafted heterotopically in the same individual, which was thus left with only the grafted organ. This was done in a very large number of cases but due to the severity of the operation only twenty-four lived.

Although mortality has been extremely high, ninety-seven animals have been obtained in the triple pituitary series and the observations upon them form the basis of this report. A small series of animals has been obtained from the second type of operation and will be briefly considered.

The results of the second type of operation, in which ectoderm only was grafted, may be dismissed by saying that they were negative. No external enlargement developed at the site of implantation as in the first type, but that the graft healed in was shown by the invagination of olfactory pits in some cases. The animals were perfectly normal in all observable characters. The material has not all been examined microscopically but in two cases in which ectoderm only was grafted no pituitary tissue was found.

Microscopic examination of those grafts in which neural tube material was taken shows pituitary tissue to be present. A discussion of the growth, differentiation and relationship of this tissue will be given at another time, as well as the results of a series of heteroplastic grafts. The present discussion will be limited to the homoplastic grafts of the first type in abnormal positions.

With one exception all the animals in which the entire pituitary region had been removed and grafted above the limb are normal as to pigmentation, growth rate and body proportions. This indicates that the grafts are functional since if they were not the animals would have had the characteristics of extirpated animals. In only one case in the twenty-four of this series was degeneration of the graft noted. This animal showed all the characteristics of a pituitaryless animal. There is the possibility of regeneration of a pituitary in the normal position. To check this, the attempt was made to remove the graft from these animals and thus to produce pituitaryless animals. Unfortunately this operation caused the death of all but five. In one of these, histological examination showed that the pituitary was completely lacking in the normal position. This case indicates that a hypophyseal transplant can function, and bring about normal development of pigment, size and body proportions in animals in which the normal pituitary is absent.

The presence of two transplanted glands in addition to their own has brought about decided changes in the pigmentation, growth rate and body proportions of these animals. The pigmentation was greatly increased and skin mounts have shown this to be due both to expansion of melanophores and to increase in their number. Of the 1928 Amblystoma punctatum, sixteen were very black, fourteen showed decided darkening, two others showed some change, and four were normal. Of the 1928 axolotls, eighteen were extremely dark, six were darker than the controls and one was normal. Of the 1929 punctatum twenty-five were extremely dark, twelve dark, and two nearly normal. The records show, then, that ninety-three per cent of the triple pituitary animals had the pigmentation increased to some degree; that in seventy-three per cent the increase was very marked and that in a large number of individual cases this increase was extreme.

In all cases the triple pituitary animal was markedly shorter than its control. This retardation of growth is not in accord with injection experiments and other similar work. The present results may be due to the continuous hypersecretion occurring during the entire period of normal gland function brought about by the addition of extra hypophyses. At the age of maximum effect, in the 1928 material, the length of the operated animals averaged about sixty-five per cent that of the control, and in the 1929 material about seventy-five per cent. There were individual cases in which the length was as low as forty per cent. The pituitaryless animals were also shorter than the controls. It will require careful analysis to determine what factors are involved in producing similar growth rates in animals which apparently have had opposite treatment.

The most striking result obtained is the disproportionate growth of the body parts in triple pituitary animals. The shortened total length

is due largely to shortening of the tail. The measurement from the tip of the snout to the anus averages about eighty-five per cent that of the control as compared to seventy-five per cent for total length. Since the tail length is extremely variable, all the measurements of body parts on both the operated animals and on the controls have been expressed in proportion to the relatively constant snout-anal length and these factors compared with each other. The tail width thus expressed is seventy-five per cent that of the control. One of the greatest changes is the shortened gills. Two sets of measurements have been made, one with, and one without the terminal filament, and the reduction is the same in each case. The gill-length is usually about seventy per cent that of the control, although there are a large number of cases in which they fall below fifty per cent. The limbs are short and thick with stubby digits, the fore limb being affected to a greater extent than the hind limb. The effect upon the proportions of the fore limb was studied by a series of measurements of its individual parts. The length of the forearm is about seventy per cent that of the control, while the width may be slightly increased, although essentially the same. This gives limbs which are disproportionately wide. Digit length is from forty to seventy per cent that of the control. The proportions of the head are also affected. The lower jaw is short and broad. There is a hypertrophy of the branchial arches which causes them to grow downward and to slip behind the opercular flap. In two cases the whole branchial basket has pushed the opercular flap in front of it. The heart is enlarged in porportion to body size. These are all easily measurable examples of the changed proportions of the animals. In the pituitaryless controls many of these changes are the reverse, there being long gills, wide tail and long thin digits. A careful analysis is being made of these effects in an attempt to distinguish between advance in differentiation and abnormality in growth.

That there is early sexual maturity is indicated by a swollen cloaca. This condition has been noted in a number of the triple pituitary animals but not in any of the controls. This observation has been substantiated in the sectioned material. A microscopic study of the gonads of both the operated and control animals is in progress to determine the extent and time of this effect.

The experimentally produced abnormalities are first observed before feeding. In extirpated animals the light pigmentation appears at Stage 41 or 42. In the grafted animals there is increase in pigmentation, retarded growth and changed body proportions at Stage 42 or 43. Stage 43 is the age at which the mouth first opens. Metamorphosis in these animals takes place normally at the same time as in the controls.

The results indicate that when a buccal rudiment is transplanted alone it will not develop, but that if neural tissue is included in the transplant

a functional gland is obtained. The excess secretion produced by transplants of the entire hypophyseal region gives retarded and disproportionate growth as well as increase in pigmentation. In Amblystoma punctatum the pituitary brings its function of pigment- and growth-regulation at Stage 42, which is a time in late embryonic life some days before feeding begins and when there is still a large amount of yolk present.

I wish to acknowledge my indebtedness to Dr. J. S. Nicholas and to Dr. Ross G. Harrison for their direction and criticism.

[1] Allen, Bennet M., *Anatom. Rec.*, **44**, 207 (1929).
[2] Harrison, Ross G., *J. Exp. Zoöl.*, **25**, 413 (1918).

THE CULTIVATION OF TREPONEMATA FROM THE BLOOD OF NORMAL MONKEYS (Macacus rhesus) AND FROM THE BLOOD OF MONKEYS INFECTED WITH YELLOW FEVER

BY ANDREW WATSON SELLARDS

DEPARTMENT OF TROPICAL MEDICINE, HARVARD MEDICAL SCHOOL*

Communicated February 17, 1930

Shortly after the demonstration of the insect transmission of yellow fever, the suggestion was made that the disease might be caused by a spirochaete. This idea was gradually received with considerable favor. Jaundice, haemorrhage into the stomach and transmission by mosquitoes are prominent features of yellow fever. Jaundice and gastric haemorrhages are striking characteristics of leptospiral infections and the spirochaete of relapsing fever is transmitted by arthropods. But the disappointment concerning leptospira in yellow fever seems to have led also to the abandonment of treponemata as a possible etiological factor since yellow fever tissues have been searched thoroughly for spiral organisms by dark field examination, by the method of Ramon y Cajal and by cultural procedures. In the following work, cultural methods have been used for the examination of monkeys infected with yellow fever.

On incubating blood from monkeys (*Macacus rhesus*) infected with yellow fever a fairly luxuriant growth of spiral organisms occurred in a number of instances. These organisms evidently belong to the genus *Treponema* according to the classification adopted by Wenyon.[1] As a general routine, infective blood was distributed in sterile test tubes, usually in quantities of 2 to 3 cc. and then incubated at 37°C., the cotton plugs being sealed with paraffin to delay evaporation. Growth of treponemata appeared in the blood of 8 monkeys usually within 3 to 4 days and also in the blood-tinged spinal fluid of one from which the blood clot itself

showed no growth. There were 17 monkeys infected with yellow fever in which the blood clots, on incubation, failed to show any treponemata. The monkeys from which the cultures were obtained were inoculated with yellow fever in various ways; 7 were injected with virulent blood, one with fluid from the anterior chamber of the eye and one was infected by the bites of mosquitoes.

Control Observations with Normal Monkeys.—Spirochaetes of the type which produce relapsing fever have been seen by several observers in direct smears of blood from monkeys.[2] These forms are clearly much larger than the treponemata found only by cultural methods.

Control cultures were made from 22 normal monkeys (*Macacus rhesus*) though obviously monkeys seriously sick with infections other than yellow fever would have afforded a stricter comparison. Blood (4 to 8 cc.) was withdrawn from the femoral vessels of 12 monkeys and 8 cc. was taken from the heart of 10 monkeys. All of the cultures remained sterile except one set from the heart blood of one monkey and in these cultures a small treponema developed after 5 days' incubation. Subcultures were maintained on clotted guinea pig blood which was first heated for 10 minutes at 55°C. Abundant growth usually occurred in 48 hours. No growth took place on blood covered with a layer of vaseline.

Stained with Giemsa, the spirals are seen to be somewhat loosely wound. In an actively growing 48-hour culture, they vary in length from about 2 to 4μ with an estimated width of about one-fifth of a micron. The average length of a complete coil was estimated at 1.8μ. Prominent beading of the spirals is of frequent occurrence. Many small clumps of spirals are usually seen in the Giemsa smear. In dark field preparations, the organisms are seen to be flexible and they show moderately rapid motion.

After subcultivation for 3 months on whole guinea pig blood, no growth occurred on guinea pig blood diluted with water (1–10) nor on blood agar (hormone agar pH 7.6 with 5% of defibrinated horse blood).

The inoculation of monkeys, intraperitoneally with large amounts of culture (7–10 cc. on guinea pig blood) produced no reaction.

Guinea pigs were not affected by injection of ordinary amounts of this culture but 3 to 5 cc. of a guinea pig blood-clot culture injected intraperitoneally produced an acute intoxication which was usually fatal within 24 hours. At autopsy, there was no jaundice, the liver appeared normal but haemorrhages occurred frequently in the subcutaneous tissue near the site of injection and in the stomach.

Cultures from Monkeys Infected with Yellow Fever.—The strains of treponemata which were isolated from yellow fever monkeys showed marked variation in their morphology. From two moribund monkeys a growth of coarse, fairly thick spirals was obtained but these strains

were lost in the early subcultures. Two strains were obtained which resembled the cultures recovered from the stock monkey.

Staining and Morphology.—The majority of the strains obtained from yellow fever monkeys showed organisms which stained a deep blue or purple with Giemsa, not altogether uniformly, but without prominent beading, and differed in detail from the treponema obtained from the normal monkey. The following details of size and measurements of the strains from yellow fever are based on Giemsa stained preparations. Many small forms are seen consisting of about one-half of a full coil and measuring 1μ or even less in length. Spirals are frequently found containing approximately 2 to $2^1/_2$ coils and measuring 3 to $3^1/_2\mu$ in length. The width is naturally difficult to determine but it was estimated to be about one-fifth of a micron. In old cultures, long forms are often found. The average length of a complete coil was estimated to be 1.2μ. The ends are tapering but no flagellae have been observed. The motility as seen in the dark field consists of rapid undulatory movements with rather slow locomotion.

Cultivation.—Subcultures of these strains, as a rule, grew out readily often within 24 to 48 hours, using guinea pig blood which had been heated 5 to 15 minutes according to the plan used by Teague for the cultivation of the Ducrey bacillus. Guinea pig blood was chosen since this animal frequently conserves the virus of yellow fever though without showing symptoms. Fair growth occurred with some strains in guinea pig blood diluted with normal saline (1–10) but not in blood diluted with water. No growth occurred on a clot of guinea pig blood covered with a layer of vaseline.

Growth at room temperature is variable but is usually much more feeble than at 37°C. After several days in the incubator, the cultures frequently break up into a mass of granules which resemble to a certain extent some of the forms seen in the stomach of infected mosquitos. A subculture from a tube showing these granules gives a luxurious growth of spiral organisms. These cultures might very well furnish an excellent starting point for applying rigid tests to determine whether the granules as seen in smears are viable. Similar masses of granules were not seen in the culture from the stock monkey. Some strains from yellow fever animals especially after prolonged subculture (9 months) in whole blood grew well on blood agar (hormone agar pH 7.6 + 5% of defibrinated horse blood) forming distinct elevated colonies, almost white, with smooth edges. Immunity tests are now being carried out with a strain growing on blood agar.

The development of cultures was definitely delayed by heating 5 minutes at 55°C., but sterilization at this temperature took place with certainty only after heating for one-half hour.

Some of these strains have remained viable for as long as six months at −7°C. The culture from the stock monkey was stored at −7°C. and was viable at the end of 3 weeks but a transplant made at the end of 2 months failed to grow.

Various procedures are now being tested for enriching the blood of monkeys for cultural purposes. The profound alterations which occur in the blood in the course of yellow fever might very well improve it as a culture medium in some respects and injure it in other ways.

Virulence.—The injection of moderate amounts of these cultures into guinea pigs and mice produced no apparent effect but cultures could be recovered from the blood of guinea pigs as long as 3 weeks after injection. The growth on 2 to 4 cc. of guinea pig blood clot when injected intraperitoneally into guinea pigs produced an intoxication which was usually fatal overnight. At autopsy, haemorrhages were frequently found in the stomach but there was no jaundice and no accumulation of fat in the liver. The majority of these strains from yellow fever animals possessed relatively little virulence for monkeys when injected in small amounts such as 1 to 2 cc. of a culture in whole blood. A single exception occurred in which the blood from an infected monkey showed very sparse growth after 8 days' incubation and the subcultures grew slowly. Morphologically, the organisms showed spirals which were often irregular and in Giemsa preparations the appearances even suggested degenerative changes. One cc. of the second subculture was estimated to contain one-millionth cubic centimeter of the original virus and these cultures had been kept 23 days in the incubator. In the majority of instances, infective blood loses its virulence when kept for 5 to 7 days at room temperature. This quantity of 1 cc. of subculture on intraperitoneal injection killed a monkey in 16 days and the liver showed entirely typical lesions of yellow fever; a growth of treponemata was obtained in blood taken just before death. One cannot strictly rule out the possibility that even after this period of 23 days a minute trace of the original virus may have persisted in the inoculum.

The remainder of the work was conducted with strains that were subcultured sufficiently long to remove any trace of the infective monkey's blood from which the cultures were obtained. Although small amounts of culture ordinarily produce no symptoms, the intraperitoneal inoculation of a normal monkey with 5 to 10 cc. of an actively growing culture in whole blood usuallypro duces profound collapse overnight, the temperature falling as much as 8°F. and death often occurs in 24 to 36 hours. At autopsy, the subcutaneous tissues are usually jaundiced, the stomach sometimes contains copious amounts of dark blood and the liver is usually very fatty. From the gross appearance, the picture is often indistinguishable from yellow fever. However, on section, the liver of the animals dying in this manner after acute intoxication showed no necrosis but only

fatty infiltration. In one instance nuclear inclusions were seen similar to those described by Torres[3] and by Cowdry and Kitchen.[4]

Pathology.—Several monkeys which received small repeated injections of culture gradually lost weight and died and in some of these the liver showed distinct lesions with necrosis. In two instances in addition to the fatty infiltration, less than half of the liver lobules remained with the normal arrangement of cells. Necrosis did not occur to the extent that usually results from injection of fully potent virus but the histological picture did resemble closely some of the less severe lesions of the liver that are occasionally seen after injection of yellow fever virus.

Immunity Tests.—Inoculation of young cultures in large amounts intraperitoneally into 3 monkeys immune to the virus of yellow fever produced no significant change in temperature. On the day after injection there was tenseness of the abdomen and some signs of discomfort but the monkeys remained well. Similar inoculations in 3 control monkeys resulted in the death of two and the third collapsed and recovered.

The active immunization of monkeys with the cultures obtained from yellow fever monkeys proved more difficult than was anticipated. It was assumed that a few injections of living culture would afford an adequate test of the immunizing power of these strains and from time to time a number of animals were tested in this manner. Seven monkeys, in all, were injected with small amounts of living culture and remained in excellent condition. Five of these proved to be susceptible to yellow fever and the other two died when injected with a large amount of living culture.

The subsequent tests were made with monkeys only after they had first been rendered immune to large doses of culture. Three strains were used for immunization. The mechanical dilution at each subculture was estimated at 1–1000. One monkey received its first inoculation with the third subculture of one strain (i.e., an estimated dilution of the original infective blood of 1–1,000,000,000). The immunization of a second monkey was started with the 4th subculture of a second strain. Subsequently these two monkeys, when given an immunity test with yellow fever virus, remained in excellent condition without losing weight or appetite, though one of the two developed fever (105^4) for one day only, on the 7th day after inoculation. These two animals are alive and well, more than 9 months having elapsed since their immunity test. Five animals were available for controls and they all died promptly.

A third strain was used for immunization which had been maintained for 34 days in culture, having been carried through 13 subcultures and then passed through a short series of normal monkeys for the purpose of increasing its virulence. An attempt was made to immunize 6 monkeys with this strain and all but two eventually succumbed to the repeated injections. The two surviving animals were injected with 1.5 cc. of

virulent blood. They showed no febrile reaction over a period of 16 days. However, one was found dead of yellow fever on the 30th day after inoculation. The other after an interval of 4 months is alive and well. Two control monkeys were injected with only 0.5 cc. of virulent blood and they died of yellow fever in 6 and 8 days, respectively.

Passive Immunity.—Another monkey (A) was immunized with this strain of the treponema from yellow fever and at the same time a control monkey (B) was immunized with the culture obtained from the stock monkey. These two animals were bled and 3 cc. of the serum of each was mixed with 0.5 cc. of virulent blood and injected intraperitoneally. A third monkey (C) received 0.5 cc. of virulent blood only; this animal was found dead on the 5th day after injection. The one injected with virus and the control serum (from monkey B) was also found dead of yellow fever on the 5th day. The one receiving serum from the monkey (A) immunized with the culture obtained from a yellow fever animal developed an abortive reaction. On the 6th day after injection the temperature rose to 104^4 and there was some loss of appetite but not of strength nor activity. On the following day the temperature was 103^8 and the animal recovered promptly. This monkey received serum in the proportion of 1.8 cc. per kilo of body weight. The remaining portion of this immune serum (1.8 cc.) was injected together with yellow fever virus into a larger monkey weighing 3 kilos (0.6 cc. of immune serum per kilo) and no protection resulted.

The development of even partial protection by immunization is regarded as significant. This work has been conducted with the French strain of yellow fever virus and it has proved very virulent for monkeys. In the course of 2 years' work, 32 normal monkeys have been injected with infective blood and all have died of yellow fever within periods varying from 4 to 12 days.

Enhancement of Virulence.—Rapid passage of these cultures through monkeys has not yet resulted in any marked increase in virulence. It is clear that the important experiments such as mosquito passage, filtrations and protection tests with human convalescent serum can be made in a convincing manner only with cultures that produce fatal infections with regularity.

Summary.—Blood cultures were made from 22 normal monkeys (*Macacus rhesus*) and growth of a treponema occurred in the blood of one animal. Similar cultures from monkeys infected with yellow fever resulted in growth of treponemata in 9 of 26 animals. The strains from the yellow fever monkeys varied greatly in their morphology. The interpretation of the observations with the strains of treponemata from the yellow fever animals is complicated by the possibility of dealing with mixed cultures of treponemata, by the tendency of spirochaetes to lose virulence in sub-

culture and by the possibility of carrying through the subcultures traces of virus from the infected animal.

Some of the strains obtained from yellow fever sources on injection in large amounts into monkeys, produced a rapidly fatal intoxication characterized by jaundice and gastric haemorrhages and a fatty liver. These features are not in themselves significant since spiral organisms wholly unrelated to yellow fever may produce extreme jaundice and extensive gastric haemorrhages. Repeated inoculation of monkeys with small amounts of some of the cultures from yellow fever animals resulted in death. On section, the liver did not show the widespread necrosis that ordinarily occurs in monkeys after the injection of yellow fever virus. However, in some instances there was extensive fatty infiltration and definite injury corresponding to the less severe lesions of the liver that occasionally occur in monkeys dying of yellow fever.

Thorough immunization of 5 monkeys with cultures of treponemata obtained from yellow fever sources resulted in partial to complete protection against the virus of yellow fever.

The treponema cultivated from the normal monkey shows differences in morphology and in its biological characteristics as compared with the majority of the strains from yellow fever monkeys. No investigation has been made to determine whether the culture obtained from the blood of a normal monkey corresponds definitely with some of the forms known to occur on the mucous membranes.

The designation *Treponema xanthogenes* appears appropriate for the organism which produces jaundice and a yellow fatty liver on injection into normal monkeys. The more fundamental characteristics of this organism are its ability to produce injury of the liver and to immunize monkeys partially or completely against the virus of yellow fever. The immunological data are sufficiently striking to require a reconsideration of treponemata in the study of the etiology of yellow fever. It is self evident that the observations are incomplete especially as regards the pathology, mosquito transmission and immunological work particularly with human convalescent serum but experimentation in these directions has proved to be very time consuming. Final interpretation must await the completion of additional experimental observations.

* It is a pleasure to acknowledge generous assistance of the DeLamar Mobile Research Fund of the Harvard Medical School in defraying the expenses of this work.

[1] Wenyon, C. M., *A Manual of Protozoölogy*, Wm. Wood & Co., N. Y., **1926**, No. 2, p. 1233.

[2] *Ibid.*, p. 1259.

[3] Torres, M., *Supp. Mem. Inst. Oswaldo Cruz.*, **1928**, No. 2, p. 55.

[4] Cowdry, E. V., and Kitchen, S. F., *Science*, **69**, 1929(252).

THE PROBLEMS OF LAGRANGE AND MAYER UNDER GENERAL END CONDITIONS

BY MARSTON MORSE

DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY

Communicated February 15, 1930

1. *Introduction.*—A general formulation of problems of this type has been given by Bolza[1] who obtained necessary conditions analogous to the Euler equations and the transversality conditions. Bliss[2] lightened Bolza's hypotheses, and gave a new form to the problem. In its simpler aspects in the plane it was recognized by Hilbert[3] and others that the conditions analogous to the Jacobi conditions could be given in terms of characteristic roots of an auxiliary boundary value problem. A necessary condition of the latter sort for the general problem has been recently obtained by Cope.[4]

Sufficient conditions in this general problem have never been established. The present note gives such conditions and outlines a proof.

This proof has been completed in detail and will be presented later in full.

The author has also studied the *complete* problem of which the minimum problem is a special case, namely, the problem of finding an extremal which gives to a certain fundamental form a prescribed type number. The solution is in terms of characteristic roots.

Finally the results obtained lead to new types of separation and comparison theorems involving the relative distribution of characteristic roots of two different auxiliary boundary value problems. This is in contrast with earlier comparison theorems,[5] more geometric in nature involving focal points and conjugate points.

2. *Necessary conditions.*—In the space of the variables x and $(y) = (y_1, \ldots, y_n)$ let there be given an ordinary[6] curve γ, $y_i = \bar{y}_i(x)$ defined on the interval $a^{(1)} \leqq x \leqq a^{(2)}$. Points neighboring the initial and final end points of γ will be denoted by $(x^s, y^s)(s = 1,2)$. With Bolza we first seek the conditions under which γ minimizes the expression

$$J = \int_{x^{(1)}}^{x^{(2)}} f(x, y, y')dx + G(x^s, y^s) \tag{2.1}$$

among curves of class C' neighboring γ and restricted as follows:

The curves shall satisfy differential equations of the form

$$\varphi_\beta(x, y, y') = 0 \quad (\beta = 1, \ldots, m) \tag{2.2}$$

and have end points given by the equations

$$y_i^s = y_i^s(\alpha) \quad x^s = x^s(\alpha) \quad (s = 1,2; i = 1, \ldots, n) \tag{2.3}$$

where (α) stands for a set of r parameters neighboring $(\alpha) = (0)$, and $(\alpha) = (0)$ corresponds to the ends of γ.

We assume that the matrix of the first partial derivatives of the functions (2.2) with respect to the variables (y') is of rank m, and that the matrix of the first partial derivatives of the $2n + 2$ functions (2.3) is r.

If γ minimizes J and is "normal," it is known that there exists m multipliers $\lambda_\beta(x)$ such that γ satisfies

$$\frac{dF}{dx} y_i' - F_{y_i} = 0 \quad F = f + \lambda_\beta \varphi_\beta. \tag{2.4}$$

For (x, y, y') on γ we set

$$\omega(\eta, \eta') = F_{y_i y_j} \eta_i \eta_i + 2F_{y_i y_j}' \eta_i \eta_j' + F_{y_i' y_j'} \eta_i' \eta_j'. \tag{2.5}$$

Suppose now we have a family of restricted curves $y_i = y_i(x, e)$ taking on end points given by (2.3) for $\alpha_h = \alpha_h(e)$. $(h = 1, \ldots, n.)$ We set $\alpha_h'(0) = u_h$ and show that for $e = 0$

$$I = \frac{d^2 J}{de^2} = b_{hk} u_h u_k + \int_{a^{(1)}}^{a^{(2)}} \omega(\eta, \eta') dx \quad (h,k = 1, \ldots, r) \tag{2.6}$$

where b_{hk} is a constant independent of the family used, and $\eta_i = y_{ie}(x, 0)$.

We also set

$$\gamma^s = \frac{\partial x^s(\alpha[e])}{\partial e} \quad e = 0, \ (s = 1, 2)$$

The variables γ^s and the end values of η_i, η_i^s, are readily seen to be connected with the variations (u) by linear relations of the form

$$\eta_i^s = c_{ih}^s u_h \ (s = 1, 2; \ i = 1, \ldots, n; \ h = 1, \ldots, r) \tag{2.7}$$

$$\gamma^s = x_{\alpha_h}^s . u_h \tag{2.8}$$

We show that (2.7) and (2.8) can be solved for the set (u) in terms of a suitable subset of r of the variations (γ^s, η_i^s). *Without loss of generality we can and will regard the variables* (u) *in* (2.6) *as a subset of* r *of the variations* (γ^s, η_i^s) *but hold to the form* (2.6) *for the sake of symmetry.*

We now replace the second variation I by an expression $I(\sigma)$ obtained by adding $\sigma \eta_i \eta_i$ to the integrand in (2.6). One is naturally led to the problem of minimizing $I(\sigma)$ subject to conditions of the following form, (see 2.7).

$$\Phi_\beta = \varphi_{\beta y_i} \eta_i + \varphi_{\beta y_i}' \eta_i' = 0 \quad \eta_i^s = c_{ih}^s u_h \tag{2.9}$$

We now set

$$\Omega(\eta, \eta', \mu, \sigma) = \omega(\eta, \eta') + \mu_\beta \Phi_\beta + \sigma \eta_i \eta_i \tag{2.10}$$

Subject to $\Phi_\beta = 0$ set

$$\zeta_i = \Omega_{\eta' i} \qquad (i = 1, \ldots, n) \tag{2.11}$$

We assume the usual bordered determinant[7] does not vanish and introduce the Hamiltonian function $H(\eta,\zeta)$. Necessary conditions in the problem of minimizing $I(\sigma)$ include the differential conditions,

$$\frac{d\eta_i}{dx} = H_{\zeta_i} \quad \frac{d\zeta_i}{dx} = -H_{\eta_i} \tag{2.12}$$

while the end conditions in (2.9) and the Bolza transversality conditions give the boundary conditions

$$\begin{gathered} \zeta_i^{(1)} c_{ih}^{(1)} - \zeta_i^{(2)} c_{ih}^{(2)} = 2b_{hk} u_k \qquad (h, k = 1, \ldots, r) \\ \eta_i^s = c_{ik}^s u_k \quad (s = 1, 2;\ i = 1, \ldots, n). \end{gathered} \tag{2.13}$$

The conditions (2.13) yield at most $2n$ linearly independent conditions on the $2n$ variables (ζ_i^s, η_i^s). The parameter σ enters in (2.12) and (2.13) only in the form $-2\sigma\eta_i$ in the extreme right member of (2.12).

Conditions (2.12) and (2.13) define the so-called *auxiliary boundary problem*. A rather different form has been obtained by Cope. The necessary condition is as follows:

If a normal extremal γ furnishes a minimum for J there can exist no characteristic roots σ in the auxiliary boundary problem for which $\sigma > 0$ and $(\eta) \not\equiv (0)$.

3. *Sufficient Conditions.*—For the general case the ordinary assumption in the one variable end point problem, that the end manifold be not tangent to the extremal, is here replaced by the following:

In the space of the $2n + 2$ variables (x^s, y_i^s) consider the 2-dimensional manifold obtained by combining an arbitrary point $(x^{(1)}, y_i^{(1)})$ on the extremal γ near its initial end-point, with an arbitrary point $(x^{(2)}, y_i^{(2)})$ on γ near its final end-point. Call this the extremal manifold.

We assume that in $(2n + 2)$-space the extremal manifold is tangent to the end-point manifold (2.3) along no regular curve. We call this the non-tangency condition. A necessary and sufficient condition for this non-tangency condition to hold is that the matrix of elements c_{ih}^s of (2.7), (2.9) and (2.13) be of rank r.

Under the non-tangency condition the set (u) appearing in (2.6) may be regarded as a subset of r of the variations η_i^s.

We prove the following theorem.

THEOREM 1. *An extremal γ, normal on every subinterval, for which the*

usual Clebsch and Weirstrass sufficient conditions hold, will afford a proper, strong relative minimum to J, under the non-tangency condition, and the condition that there exist no characteristic roots $\sigma \leqq 0$.

The proof is afforded by setting up a quadratic form in (z), $Q(z, \sigma)$ analogous to the one previously set up by the author,[8] depending for its definition upon the evaluation of $I(\sigma)$ along a finite succession of secondary extremals of $I(\sigma)$ satisfying (2.7) at the ends. It is shown that the form Q is positive definite in (z) for σ sufficiently large and positive, and is never singular except for σ equal to a characteristic root. Hence under the hypotheses of the theorem $Q(z, 0)$ is positive definite. Details complete the proof.

4. *The Complete Problem.*—A formulation of the problem of the calculus of variations which for regular problems will include all extremals under some problem is the following:

To find an extremal for which the form $Q(z, 0)$ *has a prescribed type number* k.

To seek a minimum is to require k to be zero. A theorem which gives a "solution" of this problem in the same sense that Theorem 1 gives a solution of the minimum problem is the following:

THEOREM 2. *Under the hypotheses of Theorem 1, excepting the last, the form* $Q(z, 0)$ *will have a type number equal to the number of characteristic roots* $\sigma > 0$.

Each root must be counted with a multiplicity equal to the number of corresponding independent characteristic solutions. Moreover the nullity of $Q(z, 0)$ equals the multiplicity of $\sigma = 0$.

One obtains next a large number of comparison theorems of which the following are typical.

THEOREM 3. *The number of conjugate points of* $x = a^{(1)}$ *preceding* $x = a^{(2)}$ *in the fixed end point problem, equals the number of characteristic roots greater than zero in the auxiliary boundary problem in which the boundary conditions are* $\eta_i^s = 0$.

THEOREM 4. *The number of characteristic roots* $\sigma > 0$, *in any auxiliary boundary problem for which there are* r *end parameters* (α), *lies between* k *and* $k + r$, *inclusive, where* k *is the number of characteristic roots greater than zero in the boundary problem with end conditions* $\eta_i^s = 0$.

These two theorems with a casual inspection of $I(\sigma)$ show that there are infinitely many characteristic roots[9] less than zero in any auxiliary boundary problem.

The Non-Tangency Condition Unfulfilled.—Lack of space forbids the complete summary of results for this problem. One here adds $\sigma\eta_i\eta_i$ to the integrand in (2.6) and $\sigma[\gamma^{(1)^2} + \gamma^{(2)^2}]$ outside the integral. The *extended* function $I(\sigma)$ then gives rise to an extended auxiliary problem as before. The necessary and sufficient conditions hold as before, omitting

the tangency condition, and considering characteristic roots and solutions for which either

$$\gamma^{(1)^2} + \gamma^{(2)^2} \neq 0 \quad \text{or} \quad (\eta) \not\equiv (0).$$

Further conditions peculiar to this special case are also obtained.

[1] Bolza, *Math. Ann.*, **74**, 403 (1913).

[2] Bliss, *Trans. Amer. Math. Soc.*, **19**, 305 (1918).

[3] In this connection see Lovitt, *Linear Integral Equations*, p. 207, for work of Hilbert. Richardson, *Math. Ann.*, **68**, 279 (1910). Plancherel, *Bull. Sci. Math.*, **47**, 376 (1923). Bliss, *Bull. Amer. Math. Soc.*, **32**, 317 (1926).

[4] Cope, University of Chicago Doctor's Thesis.

[5] Morse, "A Generalization of the Sturm Separation and Comparison Theorems," *Math. Annal.*, **99** (1930), about to appear.

[6] For brevity in this abstract the reader is asked to add such conditions of differentiability as are necessary.

[7] Bolza, *Vorlesungen über Variationsrechnung*, p. 589.

[8] Morse, *Trans. Amer. Math. Soc.*, **31**, 379 (1929).

[9] This is considered by Cope in his thesis in connection with the results of Bliss in the paper "A Boundary Value Problem for a System of Ordinary Linear Differential Equations of the First Order," *Trans. Amer. Math. Soc.*, **28**, 561 (1926). But although Cope has shown his boundary problem self-adjoint, the question of the *definite* self-adjointness does not seem to be settled.

CONCERNING PERFECT CONTINUOUS CURVES

BY R. L. WILDER

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN

Communicated January 29, 1930

In my paper, "Characterizations of Continuous Curves That Are Perfectly Continuous,"[1] I have shown that if M is a perfect continuous curve in a locally compact metric space, then (1) every connected subset of M is connected im kleinen, and (2) if C is a component of a G_δ in M, then C is arcwise connected. In the present paper I propose to give some applications of these results.

The advantage in considering a G_δ in M rather than simply an open subset in M will be apparent to the reader if it is noticed with what ease the general theorems stated below are obtained. This is not only true in the case of sufficiency proofs, but also in necessity proofs. For example, in the proof of Theorem 2 below, when considering the accessibility of a boundary point P of a connected set N which is a G_δ in M, it becomes evident immediately that $N + P$ is also a G_δ in M, and hence that the arcwise connectedness im kleinen enjoyed by all connected G_δ's gives at

once the regular accessibility of P from N. Of course the regular accessibility in the case of ordinary connected open subsets then follows as a special case of this result.

It is interesting to note the analogy between the quasi-open subsets (i.e., the G_δ's) of a perfect continuous curve and the open subsets of continuous curves. Every component of a G_δ in any topological space is itself a G_δ (see Lemma 2 below), so that the fact that components of open subsets of a continuous curve are themselves open[2] is paralleled, *a fortiori*. The arcwise connectedness im kleinen of the components of open subsets of a continuous curve, as well as the fact that their closed and compact subsets are imbedded in continuous curves of the component are also paralleled here. It is also found that the components of quasi-open subsets have property S[3] (indeed any connected subset has property S), and that the boundary points are regularly accessible. The last two results may seem at first glance not to contain any analogy to the case of open subsets of continuous curves, but upon closer inquiry it appears that the analogy exists; for Whyburn has shown[4] that if M is a plane continuous curve, and R a bounded, connected open subset of M, then (1) the regular accessibility of the boundary points of R and the possession of property S by R are co-existent properties, and (2) if $M - R$ is a bounded continuous curve, R has both of these properties.

That the properties of quasi-open subsets of perfect continuous curves stated in the preceding paragraph are by no means self-evident may be seen from a certain connected set obtained from a three-dimensional perfect continuous curve constructed by Whyburn.[5] Referring to a system (ρ, θ, z) of cylindrical coördinates in E_3, let AB denote the interval $(0, 1)$ of the z-axis. Let n take on, in ascending order, the set of values represented by the set of all positive prime integers. For each n, AB is the sum of a set, I_n, of n equal intervals obtained by inserting a set, K_n, of $n - 1$ points of subdivision; in the plane $\theta = \pi/n$ consider a semicircle constructed on each interval of the set I_n, having this interval as diameter, and let C_n denote the set of all these semicircles. Let $M = AB + \Sigma C_n$. Then, as shown by Whyburn, M is a perfect continuous curve, and the set N, which consists of the set ΣC_n is arcwise connected, but not arcwise connected im kleinen; also, N has a boundary point, say the point $(0, 0, {}^1/_2\sqrt{2})$, which is not accessible at all from N.

Now it will be noticed that the set $M - N$ in Whyburn's example is not an F_σ. Suppose, however, we select on AB a denumerable set, D, of points for which z is irrational, and such that D is dense in AB. Let $AB - D = C$. Then D is an F_σ, and the set $N + C$ is the only component of the quasi-open subset $M - D$. According to the results established below, not only is the set $N + C$ arcwise connected im kleinen, but any point of D is accessible, and, indeed, *regularly* accessible from $N + C$.

It will be admitted, we believe, that these facts are not intuitively evident from the example.

The following two lemmas seem to be fundamental.

LEMMA 1. *In a topological space,*[6] *the properties "K is a G_δ," "K is a G_δ relative to a closed set F" and "K is a G_δ relative to any closed set" are equivalent.*

The proof is quite evident.

LEMMA 2. *If, in a topological space, C is a component of a set K which is a G_δ, then C is itself a G_δ.*

Proof.—Let the boundary of C be denoted by B. (It is unnecessary to make a special case of the situation where B is vacuous.) Then $C + B$ is closed.

Since K is a G_δ, we have that

$$S - K = F_1 + F_2 + F_3 + \ldots,$$

where, for every n, F_n is a closed set. Let

$$B \cdot F_i = F_i^*, \ (i = 1, 2, 3, \ldots).$$

Then F_i^* is closed. For if P is a limit point of it, P belongs to F_i, since F_i is closed. Hence P is not in C, and being a limit point of boundary points of B, must be a point of B. Consequently, the set

$$B = B \cdot \Sigma F_i = F_1^* + F_2^* + F_3^* + \ldots$$

is an F_σ. Then the set C is the difference of a closed set, $C + B$, and an F_σ, B. Accordingly, C is a G_δ relative to $C + B$, and hence, by Lemma 1, a G_δ.

We may now proceed to the proofs of the main theorems.

THEOREM 1. *In a locally compact metric space, every component of a quasi-open subset of a perfect continuous curve is arcwise connected im kleinen.*

Proof.—This theorem may be proved in various ways. Perhaps the simplest proof is as follows:

Let M be a perfect continuous curve in a locally compact metric space S, and let N be a component of a quasi-open subset of M. For any point P of N, and any positive number ϵ, let $S(P, \epsilon)$ be the set of all points of S which lie at a distance from P less than ϵ. By Theorem 2 of C. P. C., N is connected im kleinen, and the component, R, of $N \cdot S(P, \epsilon)$ which contains P is a connected im kleinen set. Furthermore, R is open relative to N. Then

$$R = M - [(M - N) + (N - R)'].^7$$

The set $M - N$ is an F_σ, since by Lemma 2 N is a G_δ, and hence the set $(M - N) + (N - R)'$ is an F_σ. Accordingly, R is a G_δ. Then since by Theorem 3 of C. P. C., R is arcwise connected, the theorem is proved.

THEOREM 2. *In order that a continuum M in a locally compact metric space should be a perfect continuous curve, it is necessary and sufficient that the boundary points of the components of quasi-open subsets of M should be regularly accessible from the components.*

Proof.—The condition is necessary. For let N be a component of a quasi-open subset of M, and let P be a boundary point of N. By Lemma 2, N is itself a quasi-open subset of M. Hence $M - N$ is an F_σ. Let $M - N = K$; then K is the sum of a denumerable infinity of closed sets, $F_1, F_2, F_3, \ldots$. If F_i is any one of these sets which contains P, then $F_i - P$ is itself an F_σ, being the difference of closed sets. Consequently, $K - P$ is an F_σ, and $M - (K - P) = N + P$ is a G_δ. Accordingly, by Theorem 1, $N + P$ is arcwise connected im kleinen at P, which proves the condition necessary.

The condition is sufficient. By proceeding as in the proof of the sufficiency of Theorem 3 of C. P. C., we obtain the component Q of $W - K$ containing all of the sets $M_1, M_2, M_3, \ldots$. It is clear that Q is a G_δ in W, and hence in M. However, if P is a point of M_ω in I, then P is a boundary point of Q, and as there are points of the sets M_i arbitrarily near to P, it follows that P is not regularly accessible from Q.

Theorem 2 is a generalization of a theorem of Whyburn[4] to the effect that if R is a component of an open subset of a perfect continuous curve in the plane, then every point of the boundary of M is regularly accessible from R.

It will now be shown that those connected subsets of perfect continuous curves that are compact have property S, which, as shown by R. L. Moore, is a stronger property than that of connectivity im kleinen.[3] In this connection it should be noted that Whyburn has shown[5] that those bounded plane continua that are perfect continuous curves are characterized by the fact that their connected open subsets have property S.

THEOREM 3. *Every compact connected subset of a perfect continuous curve in a locally compact metric space has property S.*

Proof.[9]—Let N be a compact connected subset of a perfect continuous curve M, and suppose that N does not have property S. Then there exists a positive number ϵ such that N is not the sum of a finite number of connected sets of diameter less than ϵ.

For every point P of N let $N(P)$ denote the set of all points of N that lie, with P, in a connected subset of N whose diameter is less than $\epsilon/2$. Then N contains an infinite sequence of points, $P_1, P_2, P_3, \ldots$, having a sequential limit point P, and such that no two points P_k and P_j $(k \neq j)$ of the sequence can be joined by a connected subset of N of diameter less than $\epsilon/2$. The set $N + P$, being connected, is connected im kleinen by Theorem 2 of C. P. C.

Let $S(P, \epsilon/4)$ and $S(P, \eta)$ be compact spherical neighborhoods of P,

such that $\eta < \epsilon/4$, and every point of N in $S(P, \eta)$ is joined to P be a connected subset of $N + P$ that lies in $S(P, \epsilon/4)$. There exists a number m, such that for $n > m$, P_n lies in $S(P, \eta)$. We may suppose that ϵ is so small that N contains points of F, the boundary of $S(P, \epsilon/4)$.

By Theorem 2 of C. P. C., N is connected im kleinen, and hence by the theorem in my paper, "A Theorem on Connected Point Sets Which Are Connected im Kleinen," [10] there exists, for every $n > m$, a set H_n which is a subset of N and of $S(P, \epsilon/4)$, contains P_n, and has at least one limit point in $F \cdot N$. It is clear that no two sets H_k and H_j $(k \neq j)$ have a point in common.

Let the component of the set $(N + P) \cdot S(P, \epsilon/4)$ determined by P be denoted by R. Then R is connected im kleinen and by the theorem just referred to above R contains, for every $n > m$, a connected set K_n which has both P_n and P as limit points, but contains neither of these points. The set

$$T_n = H_n + K_n$$

is a connected subset of N having both P and points of $F \cdot N$ as limit points, and lying wholly in $S(P, \epsilon/4)$. Since T_n contains P_n, it is clear that no two sets of type T_n have a point in common, and that no connected subset of N which lies wholly in $S(P, \epsilon/4)$ contains points of two such sets. We shall suppose, then, that each set T_n is the component of the set $N \cdot S(P, \epsilon/4)$ determined by P_n.

Denote the set $N \cdot S(P, \epsilon/4)$ by H. Then, since H is connected im kleinen, the components T_n are also quasi-components of H. Consequently, for any positive integer $k > m$, we can say

$$H = T_k + H_1 \text{ separate},$$

where H_1 is the set $H - T_k$.

Now let ρ be a positive number such that $\epsilon/4 > \rho > \eta$ and for some fixed $k > m$ let

$$C_1 = S'(P, \rho) \cdot T_k' \cdot H'_1.$$

The set C_1 so defined contains no point of H, and no connected subset of $N' - C_1$ that lies in $S'(P, \rho)$ contains both P_k and P_{k+j} $(j > 0)$, as can be shown by the same methods employed in a similar situation in the proof of Theorem 2 of C. P. C.

Since N' is a continuous curve and N contains no point of C_1, there exists, by virtue of a theorem of R. L. Moore,[11] an arc in $N' - C_1$ joining P_k to a point Q of N which is not in $S(P, \rho)$. Let Q_k be the first point of B, the boundary of $S(P, \rho)$, on this arc in the order from P_k to Q.

Again, we have

$$H = T_{k+1} + H_2 \text{ separate}$$

where $H_2 = H - T_{k+1}$, and a set C_2 defined as follows:

$$C_2 = S'(P, \rho) \cdot T_{k+1}' \cdot H_2'.$$

It can be shown, as above, that $N' - (C_1 + C_2)$ contains an arc $P_{k+1}Q_{k+1}$ which joins P_{k+1} to a point Q_{k+1} of B, and lies, except for Q_{k+1}, wholly in $S(P, \rho)$. It is easy to see that the arcs P_kQ_k and $P_{k+1}Q_{k+1}$ have no point in common.

Proceeding in this way indefinitely, and letting R_1 denote the component of the set $(N + P) \cdot S(P, \eta)$ determined by P, we have a connected set

$$A = \sum_{j=r}^{\infty} P_{k+j}Q_{k+j} + R_1,$$

where r denotes that the point P_{k+r} is the first point of the sequence $P_k, P_{k+1}, P_{k+2}, \ldots$, which lies in R_1. The set A' is a subset of N' and hence a continuous curve. However, by a method similar to that used in the last two paragraphs of the proof of Theorem 2 of C. P. C., it can be shown that A' is not connected im kleinen at some point whose distance from P is less than ρ but greater than η.

I have shown elsewhere[12] that if K is a closed and bounded set of points lying in a component Q of an open subset of a continuous curve M (where M may be assumed to lie in any locally compact metric space), then K is a subset of a continuous curve N which itself lies in Q. The following theorem is the analogous theorem for perfect continuous curves.

THEOREM 4. *If K is a closed and compact set of points lying in a component Q of a quasi-open subset of a perfect continuous curve M which is imbedded in a locally compact metric space, then K is a subset of a continuous curve N which itself lies in Q. The curve N is also a curve in the Menger sense, and such that $N - K$ is the sum of a denumerable infinity of mutually exclusive open arcs only a finite number of which are of diameter greater than any given positive number.*

Proof.—In my paper, "On the Imbedding of Subsets of a Metric Space in Jordan Continua" (not yet published), I have shown that in order that a compact and closed subset H of a metric space R should be a subset of a continuous curve J in R, it is necessary and sufficient that H should be arcwise connected through R and arcwise connected im kleinen through R, and that, moreover, if these conditions are satisfied, J may be so chosen that the complement of H (relative to J) is a denumerable set of mutually exclusive open arcs whose diameters converge to zero, and such that if M is of positive dimension h, the dimension of J is likewise h. By virtue of Theorem 3 of C. P. C., and Theorem 1 above, the conditions stated in this result are satisfied for the set K and the metric space Q. That the curve N stated in the theorem is a Menger curve follows from a result

of J. R. Kline,[13] to the effect that every perfect continuous curve is of dimension one in the Menger-Urysohn sense.

In my paper, "A Property Which Characterizes Continuous Curves,"[14] I have shown that those continua which are continuous curves are characterized by a property which I called *normal connectivity*. To recall this property: If M is a continuum and C_1 and C_2 are mutually exclusive closed subsets of M, we define a set $K(C_1, C_2)M$ as the sum of a connected subset H of M which has limit points in both C_1 and C_2, but contains no point of $C_1 + C_2$, and the set of all limit points of H in $C_1 + C_2$. Then M is normally connected if every connected subset which contains points of two such mutually exclusively closed subsets, C_1 and C_2, also contains at least one point of a set $K(C_1, C_2)M$. It may be of interest to note here how this characterization of continuous curves may be modified so as to characterize perfect continuous curves. Let us call a set of the type of H above, a set $H(C_1, C_2)M$. Then we can state the following theorem:

THEOREM 5. *In order that a continuum M in a locally compact metric space should be a perfect continuous curve it is necessary and sufficient if C_1 and C_2 are any two mutually exclusive closed subsets of M, and N is a connected subset of M containing points of both C_1 and C_2, that N contain a set $H(C_1, C_2)M$.*

Proof.—That the condition stated is necessary follows at once from Theorem 2 of C. P. C. and the theorem stated in my paper, "A Theorem on Connected Point Sets Which Are Connected im Kleinen."[10]

The condition is sufficient. Let M be a continuum satisfying the conditions of the theorem, and suppose U is a subcontinuum of M. Let C_1 and C_2 be mutually exclusive closed subsets of U, and N a connected subset of U containing points of both C_1 and C_2. By hypothesis, N contains a set $H(C_1, C_2)M$, which is, *a fortiori*, a set $H(C_1, C_2)U$. Consequently, since this set $H(C_1, C_2)U$, together with its limit points in $C_1 + C_2$, is a set $K(C_1, C_2)U$, the set U is normally connected, and therefore a continuous curve.

In conclusion, we may also note that as a result of Theorem 2 of C. P. C. and Theorems 1 and 2 of my paper, "A Property Which Characterizes Continuous Curves," perfect continuous curves are characterized by the fact that *their connected subsets are normally connected.*

[1] These PROCEEDINGS, **15**, 1929 (614–621). This paper will be referred to hereinafter as C. P. C. A continuous curve is said to be perfect if every subcontinuum of it is itself a continuous curve.

[2] Kuratowski, C., *Fund. Math.*, **1**, 1920 (40–43).

[3] A connected set M is said to have property S if for every $\epsilon > 0$ M is the sum of a finite number of connected sets of diameter less than ϵ. Cf. Moore, R. L., *Fund. Math.*, **3**, 1922 (232–237).

[4] Whyburn, G. T. these PROCEEDINGS, **9**, 1927 (650–657).

[5] Cf. abstract, *Bull. Amer. Math. Soc.*, **34**, 1928 (551). Whyburn's paper, "Concerning Points of Continuous Curves Defined by Certain im Kleinen Properties," which contains this example in full, is to appear in an early issue of *Math. Ann.* My ability to reproduce it here is due to Professor Whyburn's kindness in communicating it to me.

[6] Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, **1914**, p. 213.

[7] If K is a set of points, by K' we denote the set consisting of K together with its limit points.

[8] These PROCEEDINGS, **12**, 1926 (761–767).

[9] Compare the first part of this proof with that given by Whyburn for the case cited above.

[10] Wilder, R. L., *Bull. Amer. Math. Soc.*, **32**, 1926 (338–340). It is clear from its proof that this theorem is true in any topological space.

[11] Moore, R. L., *Math. Zs.*, **15**, 1922 (254–260).

[12] Wilder, R. L., *Bull. Amer. Math. Soc.*, **34**, 1928 (649–655). The theorem quoted here was proved for euclidean spaces, but it is clear that it holds in any locally compact metric space, and, indeed, may be proved directly by use of the imbedding theorem employed in proving Theorem 4 below, and hence so that N is of the same dimension (Menger-Urysohn) as H, etc.

[13] Kline, J. R., abstract, *Bull. Amer. Math. Soc.*, **34**, 1928 (263).

[14] Wilder, R. L., these PROCEEDINGS, **11**, 1925 (725–728). Although the proof is given in the terminology of the plane, it is clear that the theorem referred to is true in any locally compact metric space; in fact, the necessity proof holds in any topological space.

COMBINATORY TOPOLOGY OF CONVEX REGIONS

BY M. H. A. NEWMAN

ST. JOHN'S COLLEGE, CAMBRIDGE, ENGLAND

Communicated January 27, 1930

In a recent paper[1] I have shown that a set of rectilinear simplexes filling a convex region in R_n (Euclidean n-space) form an E_n, i.e., an n-element in the combinatory sense. The first part of the proof there given may be replaced by the considerably shorter argument which follows. For the idea of this simplification I am indebted to Professor J. W. Alexander.

The theorem to be proved is as follows:

If A_n is a finite collection of (closed) rectilinear n-simplexes in R_n satisfying the conditions: (R_1) any two simplexes have either nothing or a k-component in common ($O \leqq k \leqq n - 1$); (R_2) the points of A_n form a convex region in R_n; then the set of simplexes form an E_n.

Consider a convex region divided into convex rectilinear n-cells of any type, where by a convex rectilinear n-cell is meant a closed bounded region cut out of R_{n+1} by a finite number of flat $(n - 1)$-folds. A set of simplexes satisfying R_1 and R_2 can be obtained from such a collection of

cells by regular subdivision, i.e., by replacing the 1-, 2-, ..., n-cells successively by 1-, 2-, ..., n-stars of simplexes. An argument used in the paper referred to[2] shows that if it could be proved that regular subdivision of any such convex collection of convex cells (in particular of A_n), leads to an E_n, it would follow that A_n itself is an E_n. We wish then to prove

THEOREM A. *If K_n is a set of convex rectilinear n-cells filling a convex region in R_n, in such a way that the common part of two cells is either nothing or a k-cell of both ($O \leq k \leq n-1$), then the set of simplexes obtained by regular subdivision of K_n is an E_n.*

The theorem being trivial if $n = 1$, suppose it true of sets of cells in R_{n-1}.

We can derive from K_n another collection of cells K_n^*, of a more special type, by "producing" all the $(n-1)$-dimensional faces, i.e., introducing as barriers the complete intersection of K_n with all flat $(n-1)$-folds containing an $(n-1)$-face of an n-cell. This has the effect of replacing each n-cell of K_n by a collection of n-cells of the type of K_n itself. By the argument already referred to[2] if Theorem A can be proved for K_n^* its truth follows for K_n.

K_n^* has these properties: (P_1) no two boundary $(n-1)$-cells of the same n-cell lie in the same flat $(n-1)$-fold, (P_2) if a flat $(n-1)$-fold contains an $(n-1)$-cell all its points in K_n^* belong to $(n-1)$-cells.

(*a*) Suppose K_n^* consists of a single cell. Let C_{n-1} be any boundary $(n-1)$-cell, Π_{n-1} the sum of the remaining $(n-1)$-cells,[3] π_{n-1} the $(n-1)$-fold containing C_{n-1}. Let V be a point separated from the interior of K_n^* by π_{n-1}, but so near the center of gravity of C_{n-1} that if a straight line through V meets K_n^* it meets C_{n-1}. Since K_n^* is convex it follows that every line through V which contains a point of K_n^* contains one point of C_{n-1} and one point of Π_{n-1}. Thus projection of Π_{n-1} on to π_{n-1} from V gives a (1,1)-continuous representation of Π_{n-1} as a set of convex rectilinear $(n-1)$-cells filling the convex region C_{n-1}. The conditions of Theorem A are satisfied (with $n-1$ for n) and Π_{n-1}, on regular subdivision, becomes an E_{n-1}. Hence, by a known theorem,[4] $C_{n-1} + \Pi_{n-1}$, the complete boundary of K_n^*, becomes on regular subdivision a Σ_{n-1}—a combinatory $(n-1)$-sphere. It now follows[5] that regular subdivision of K_n^*—the join of a central vertex to a Σ_{n-1}—gives an E_n.

(*b*) The proof for any set of cells K_n^* satisfying (P_1) and (P_2) now proceeds by induction through the number of cells. Let ρ_{n-1} be any $(n-1)$-fold containing an $(n-1)$-cell interior to K_n^*. No cell of K_n^* is divided in two by ρ_{n-1}: two cells lying on its two sides form two collections of n-cells satisfying (P_1) and (P_2), which by an inductive hypothesis, may be assumed to become two E_n's on regular subdivision. Their common points form a set of $(n-1)$-cells, satisfying (P_1) and (P_2), i.e., on regular subdivision they form an E_{n-1}. Hence[6] the sum of the two E_n's is itself an E_n.

This completes the proof of Theorem A.

[1] *Proc. Lond. Math. Soc.*, **30**, 1930 (339–346). For definitions of combinatory entities see two papers, *Proc. Roy. Soc. Amsterdam*, **29**, 1926 (610–641), here cited as *F* I and *F* II.

[2] Loc. cit., p. 343, par. (*d*).

[3] All simplexes and cells in this paper are closed sets.

[4] *F* II, Theorem 6, Coroll. 1.

[5] *F* I, 12.

[6] *F* II, Theorem 8*a*, Coroll.

SOME REMARKS ON THE PROBLEM OF PLATEAU

By Tibor Radó

Department of Mathematics, Harvard University

Communicated December 19, 1929

1. The following remarks are concerned with a recent paper of R. Garnier on the problem of Plateau.[1] This problem, as investigated by Garnier according to the ideas of Riemann, Weierstrass and Schwartz, may be stated in the following analytic form:

Given, in the xyz-space, a simple closed curve Γ^*. Determine three functions $x(u,v)$, $y(u, v)$, $z(u, v)$ with the following properties.

I. $x(u, v)$, $y(u, v)$, $z(u, v)$ are harmonic for $u^2 + v^2 < 1$

II. Satisfy for $u^2 + v^2 < 1$ the relations

$$x_u^2 + y_u^2 + z_u^2 = x_v^2 + y_v^2 + z_v^2,$$
$$x_u x_v + y_u y_v + z_u z_v = 0.$$

III. $x(u, v)$, $y(u, v)$, $z(u, v)$ remain continuous on the unit circle $u^2 + v^2 = 1$ and carry the unit circle in a one-to-one and continuous way into the given curve Γ^*.[2]

We shall see (§11), that at least for a certain class of curves Γ^* the surface $x = x(u, v)$, $y = (u, v)$, $z = z(u, v)$ is uniquely determined by the above conditions; therefore it is not advisable to restrict the problem by any further conditions.

2. The solution, given by Garnier, of the problem of Plateau, can be greatly simplified by the following

Theorem of Approximation.—Let Γ_n^* be a sequence of simple closed curves in the xyz-space with the following properties.

(α) The length $l(\Gamma_n^*)$ of Γ_n^* is uniformly bounded: $l(\Gamma_n^*) \leqq L$.

(β) The problem of Plateau (as stated above) is solvable for Γ_n^* ($n = 1, 2, \ldots$).

(γ) The sequence Γ_n^* converges in the sense of Fréchet toward a simple closed curve Γ^*.[3]

Then the problem of Plateau is also solvable for the limit curve Γ^*.[4] This theorem can be applied to the work of Garnier as follows. Garnier solves the problem of Plateau in the first place for simple closed *polygons*, by solving the monodromic group problem of Riemann under very difficult supplementary conditions. This being done, he takes up the case of a simple closed curve Γ^*, approximates it by inscribed polygons Π_n^*, and investigates the solution corresponding to Π_n^* for $n \longrightarrow \infty$. He does this by means of a discussion of the solution of a monodromic group problem with variable initial conditions.[5] This complicated discussion can be avoided by the above theorem of approximation. Indeed, the lengths of the inscribed polygons are bounded by the length of Γ^* and the inscribed polygons converge toward Γ^* precisely in the sense of Fréchet; so the theorem of approximation yields immediately the result that *if the problem of Plateau is solvable for any simple closed polygon, then it is solvable for any simple closed rectifiable curve.*[6]

The method of Garnier for the case of a polygon is a *method of continuity* (Kontinuitätsbeweis). The reader will find several opportunities to simplify this part of the work of Garnier by the use of our theorem of approximation.

3. The method of Garnier is not appropriate for the discussion of the question of the *uniqueness;* even for the case of a skew quadrilateral he admits the possibility of the existence of several solutions,[7] in contradiction to a statement of Schwartz.[8] The theorem of uniqueness, proved in §11 may therefore be of some interest in spite of its very particular character.

4. *Lemma.*—Let there be given n functions $h_1(u, v), \ldots, h_n(u, v)$, continuous for $u^2 + v^2 \leqq 1$ and harmonic for $u^2 + v^2 < 1$. Let $l(\rho)$ $(0 < \rho \leqq 1)$ denote the length of the curve

$x_1 = h_1(\rho \cos\theta, \rho \sin\theta), \ldots, x_n = h_n(\rho \cos\theta, \rho \sin\theta), 0 \leqq \theta \leqq 2\pi,$

$(l(1)$ may be infinite). *Then* $l(\rho)$ *is a monotonically increasing function of* ρ *for* $0 < \zeta \leqq 1$. *In particular, if* $l(1)$ *is finite and* $\leqq L$, *then* $l(\rho) \leqq L$ *for* $0 < \rho \leqq 1$.

If $n = 2$ and if h_1 and h_2 are *conjugate* harmonic functions, this statement reduces to a nice theorem concerning conformal mapping, proved independently by P. Csillag and L. Bieberbach. The proof of Bieberbach[9] extends immediately to our general statement; the easy and attractive details may be left to the reader.

5. In order to prove the theorem of approximation, take three distinct points A, B, C on the unit circle $u^2 + v^2 = 1$ and three distinct points, A^* B^*, C^* on Γ^*; then, by virtue of condition γ (§2), it is possible to take three distinct points A_n^*, B_n^*, C_n^* on Γ_n^* converging toward A^*, B^*, C^*, respectively. The problem of Plateau being solvable by hypothesis for Γ_n^*, we have three harmonic functions $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ with

the properties listed in the statement of the problem; for the sake of clarity, we denote the boundary values of $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ by $\xi_n(\theta)$, $\eta_n(\theta)$, $\zeta_n(\theta)$, θ being the polar angle. Applying a one-to-one and conformal mapping of the unit circle upon itself (with preserved or reversed indicatrix), we can suppose that we have

$$\begin{aligned}(\xi_n(A), \eta_n(A), \zeta_n(A)) &= A_n^*,\\ (\xi_n(B), \eta_n(B), \zeta_n(B)) &= B_n^*,\\ (\xi_n(C), \eta_n(C), \zeta_n(C)) &= C_n^*.\end{aligned} \tag{1}$$

The equations $x = \xi_n(\theta)$, $y = \eta_n(\theta)$, $z = \zeta_n(\theta)$ carry the unit circle in a one-to-one and continuous way into the curve Γ_n^*; from α and γ (§2) it follows then that the sequences $\xi_n(\theta)$, $\eta_n(\theta)$, $\zeta_n(\theta)$ are uniformly bounded and of uniformly bounded total variation. By virtue of a well-known theorem of Helly, we can pick out from them *everywhere* convergent subsequences—in order to avoid complicated notations, we denote these subsequences by $\xi_n(\theta)$, $\eta_n(\theta)$, $\zeta_n(\theta)$ again. The limit functions $\xi(\theta)$, $\eta(\theta)$, $\zeta(\theta)$ are bounded and of bounded variation. Denote now by $x(u, v)$, $y(u, v)$, $z(u, v)$ the harmonic functions determined by the Poisson integral formula, $\xi(\theta)$, $\eta(\theta)$, $\zeta(\theta)$ being the boundary functions. *We shall show that these harmonic functions solve the problem for the limit curve* Γ^*.

6. Uniformly bounded sequences being integrable term-by-term, $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ converge in the interior of the unit circle toward $x(u, v)$, $y(u, v)$, $z(u, v)$, the convergence being uniform on any closed point-set interior to the unit circle; moreover, the same holds for all of the partial derivatives. From this latter remark, it follows immediately that $x(u, v)$, $y(u, v)$, $z(u, v)$ satisfy the condition II (§1). From the same remark we infer that

$$l(\rho) = \lim l_n(\rho) \text{ for } 0 < \rho < 1, \tag{2}$$

where $l(\rho)$, $l_n(\rho)$ denote the lengths of the images of the circle $u^2 + v^2 = \zeta^2 < 1$ by means of x, y, z and x_n, y_n, z_n, respectively.

Apply now the lemma of §4. With regard to condition α (§2) it follows that $l_n(\rho) \leqq L$, and therefore we get from (2)

$$l(\rho) = \int_0^{2\pi} (x_\theta^2 + y_\theta^2 + z_\theta^2)^{1/2} d\theta \leqq L; \tag{3}$$

by using the relations

$$\begin{aligned}x_u^2 + y_u^2 + z_u^2 &= x_v^2 + y_v^2 + z_v^2,\\ x_u x_v + y_u y_v + z_u z_v &= 0\end{aligned} \tag{4}$$

we obtain further from (3)

$$\rho \int_0^{2\pi} (x_u^2 + y_u^2 + z_u^2)^{1/2} d\theta = \rho \int_0^{2\pi} (x_v^2 + y_v^2 + z_v^2)^{1/2} d\theta \leqq L, \qquad (5)$$

the integrations being extended over any circle $u^2 + v^2 = \rho^2 < 1$. Now we can write

$$x(u, v) = \Re f_1(w),\ y(u, v) = \Re f_2(w),\ z(u, v) = \Re f_3(w), \qquad (6)$$

where $f_1(w)$, $f_2(w)$, $f_3(w)$ are analytic functions of $w = u + iv$ ($|w| < 1$). Denote by $T_k(\rho)$ the total variation of $f_k(\rho)$ on the circle $|w| = \rho < 1$ ($k = 1, 2, 3$). We have

$$T_1(\rho) = \int_0^{2\pi} |f_1'|\,|dw| = \int_0^{2\pi} (x_u^2 + x_v^2)^{1/2}\,|dw|,$$

etc.; from (5) we readily get that

$$T_k(\rho) \leqq 2L \text{ for } 0 < \rho < 1,\ k = 1, 2, 3. \qquad (7)$$

The total variation of $f_k(w)$ on concentric circles being bounded, $f_k(w)$ remains continuous for $|w| = 1$, by virtue of a theorem of F. Riesz;[10] with regard to (6) *it follows from this that* $x(u, v)$, $y(u, v)$, $z(u, v)$ *remain continuous for* $u^2 + v^2 = 1$. Denote their boundary values by $x(\theta)$, $y(\theta)$, $z(\theta)$.

7. The functions $\xi(\theta)$, $\eta(\theta)$, $\zeta(\theta)$ are of bounded variation; the set G of their common points of continuity is therefore certainly *everywhere dense* on the unit circle. From a well-known property of the Poisson integral formula it follows that

$$x(\theta) = \xi(\theta),\ y(\theta) = \eta(\theta),\ z(\theta) = \zeta(\theta) \text{ on } G. \qquad (8)$$

8. P being any point on the unit circle, put

$$\begin{aligned} \overline{P}^* &= (x(P), y(P), z(P), \\ P^* &= (\xi(P), \eta(P), \zeta(P), \\ P_n^* &= (\xi_n(P), \eta_n(P), \zeta_n(P)). \end{aligned} \qquad (9)$$

If P describes the unit circle, P_n^* describes just once the curve Γ_n^*. This curve converging in the sense of Fréchet toward Γ^*, and P^* being the limit of P_n^*, the following facts are immediate.[11]

I. P^* is on Γ^*; in particular the equations $x = \xi(\theta)$, $y = \eta(\theta)$, $z = \zeta(\theta)$ carry the three fixed points A, B, C into the three fixed points A^*, B^*, C^* (see (1)).

II. Let P and Q be any two points on the unit circle. Denote by σ, τ the two closed arcs bounded by them on the unit circle. The corresponding points P^*, Q^* on Γ^*, defined by (9), divide Γ^* into two closed arcs (if P^*, Q^* coincide, one of them reduces to a point). *It is possible*

to denote these two arcs by σ^*, τ^* so that the image S^* of any point S of σ be a point of σ^*, and the image T^* of any point T of τ be a point of τ^*.

III. *In particular, if P^*, Q^* coincide, then $\xi(\theta)$, $\eta(\theta)$, $\zeta(\theta)$ reduce all three to constants on one of the arcs σ, τ.*

9. Consider now a *small* arc σ on the unit circle, *bounded by two points P, Q of G;* denote by τ the complementary arc of the unit circle, and by σ^*, τ^* the two arcs of Γ^* defined by II, §8. These latter arcs are bounded by the points $P^* = \overline{P}^*$, $Q^* = \overline{Q}^*$ (see §7), defined by (9); the functions $x(\theta)$, $y(\theta)$, $z(\theta)$ being continuous and σ being a small arc, the distance of P^* and Q^* is small. Γ^* being a *simple* continuous curve, it results from this that the diameter of one of the two arcs σ^*, τ^* is small. Now σ being a small arc, τ will contain *at least two* of the three fixed points A, B, C and therefore τ^* contains at least two of the three fixed distinct points A^*, B^*, C^*. The diameter of τ^* cannot be therefore less than a fixed number, and so it results that σ^* is a small arc. *That means that if P describes a small arc σ of the unit circle, the corresponding point P^* moves on a small arc σ^* of Γ^*, provided that σ be bounded by two points of G. The set G being everywhere dense it results from this that $\xi(\theta)$, $\eta(\theta)$, $\zeta(\theta)$ are continuous on the whole unit circle $u^2 + v^2 = 1$.* We have, therefore, on the whole unit circle,

$$x(\theta) = \xi(\theta),\ y(\theta) = \eta(\theta),\ z(\theta) = \zeta(\theta). \tag{10}$$

10. From (10) and I, §8, it results that if P describes the unit circle, the point $\overline{P}^* = (x(P), y(P), z(P))$ moves on Γ^* in a continuous way. Suppose now that two distinct points P, Q have the same image on Γ^*; from (10) and III, §8, it results that on a whole arc σ of the unit circle $x(\theta)$, $y(\theta)$, $z(\theta)$ reduce all three to constants. Then $x(u, v)$, $y(u, v)$, $z(u, v)$ are certainly analytic on this arc σ and we have

$$\left.\begin{aligned} x_\theta &= -x_u \sin\theta + x_v \cos\theta = 0, \\ y_\theta &= -y_u \sin\theta + y_v \cos\theta = 0, \\ z_\theta &= -z_u \sin\theta + z_v \cos\theta = 0 \end{aligned}\right\} \text{ on } \sigma.$$

Square, add, and use (4); it results that x_u, x_v, y_u, y_v, z_u, z_v vanish on σ and that consequently $x(u, v)$, $y(u, v)$, $z(u, v)$ reduce to constants *identically*. This conclusion is, however, contradictory to the fact, appearing from (10) and I, §8, that the equations $x = x(\theta)$, $y = y(\theta)$, $z = z(\theta)$ carry the points A, B, C into three *distinct* points A^*, B^*, C^*.

The proof of the approximation theorem is thus finished. The arguments used would lead easily to the more precise result that the solution $x_n(u, v)$, $y_n(u, v)$, $z_n(u, v)$ of the problem for Γ_n^* being normalized as indicated by (1), it is possible to pick out a subsequence converging toward a solution of the problem for Γ^* *uniformly* in the closed unit circle $u^2 + v^2 \leqq 1$.

11. *A theorem of uniqueness.*[12] Suppose that the orthogonal projection of the given curve Γ^* upon some plane is a *simply covered convex curve* Γ'. Let $x(u, v)$, $y(u, v)$, $z(u, v)$ be a solution of the problem of Plateau for Γ^*. *Then the surface* $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$ *is a one-to-one image of* $u^2 + v^2 \leqq 1$ *and it is uniquely determined.*

Proof.—We can suppose that Γ' is situated in the xy-plane. The equations

$$x = x(u, v), \quad y = y(u, v) \tag{11}$$

carry then the unit circle $u^2 + v^2 = 1$ in a one-to-one and continuous way into the *convex* curve Γ'; $x(u, v)$, $y(u, v)$ being harmonic, it results from this[13] that

(*a*) The equations (11) define a one-to-one and continuous transformation of the whole region $u^2 + v^2 \leqq 1$ into the closed region bounded by Γ';

(*b*) The determinant $x_u y_v - x_v y_u$ is different from zero for $u^2 + v^2 < 1$.

That means that from (11) we obtain u and v as single-valued and continuous functions in the closed region bounded by Γ', and analytic inside of Γ'. $z(u, v)$ becomes in this way a function of u, v with the same properties and the first half of the statement is proved; this function is besides a solution of the partial differential equation

$$(1 + q^2)r - 2pqs + (1 + p^2)t = 0, \tag{12}$$

and its boundary values on Γ' are univocally determined by the mutual position of Γ^* and Γ'. The solution of the boundary value problem for (12) being unique, the second half of the statement is also proved.

[1] R. Garnier, "Le problème de Plateau," *Ann. sc. école normale supér.*, 1928, pp. 53–144.

[2] The geometric meaning of these conditions is as follows. III means that the surface $x = x(u,v)$, $y = y(u,v)$, $z = z(u,v)$ be bounded by the given curve Γ^*; II means that u^5v be isothermic parameters and then I secures by virtue of a theorem of Weierstrass, that the surface be a minimal surface.

[3] That is to say, the *distance* of Γ^* and Γ_n^* converges toward zero. The *distance* of two simple closed curves C', C'', in the sense of Fréchet, may be defined as follows. Denote by T a one-to-one and continuous correspondence between C' and C'', and by $M(T)$ the maximum of the distance of corresponding points. The lower limit of $M(T)$ for all possible correspondences T is the distance of C' and C'' in the sense of Fréchet.

[4] The situation is analogous to that in the theory of the conformal mappings of plane regions; see the author's paper *Sur les représentations conformes de domaines variables* (Acta Szeged, 1), and the papers of Courant referred to there, and the paper of Bieberbach cited in reference 9).

[5] Loc. cit., reference 1, pp. 116–144.

[6] This result is more general than the result of Garnier, whose method is restricted to curves with *bounded curvature*. The title of a paper of Dr. J. Douglas (Solution of the Problem of Plateau When the Contour Is an Arbitrary Jordan Curve in n-Dimensional Euclidian Space, to be read by title at the meeting of the American Math. Society, in Bethlehem, Dec. 27, 1929) announces a much more general result. Our point is, however, the gain in clarity rather than the gain in generality.

[7] Loc. cit., reference 1, pp. 106–107.

[8] H. A. Schwartz, "Gesammelte math. Abhandlungen," **1**, pp. 111, note (7).

[9] L. Bieberbach, "Über die konforme Kreisabbildung nahezu kreisförmiger Bereiche," *Sitzungsber. preuss. Akadie. Wiss., Math.-Phys. Klasse*, 1924, pp. 181–188.

[10] F. Riesz, "Über die Randwerte einer analytischen Funktion," *Math. Zeitschrift*, **8**, 1923, p. 95.

[11] Cf. Kerékjártó, "Vorlesungen über Topologie," passim.

[12] I found this theorem in trying to verify the statement of Schwartz referred to in § 3; I developed its proof, in a more general form, in a talk in the math. colloquium in Leipzig, June, 1925.

[13] T. Radó, *Aufgabe*, **41**, and H. Kneser, *Lösung der Aufgabe*, **41**, *Jahresbericht der deutschen Mathematiker-Vereinigung*, **35**, 1926.

THE PROBLEM OF THE CALCULUS OF VARIATIONS IN m-SPACE WITH END-POINTS VARIABLE ON TWO MANIFOLDS

BY ELDRED CURRIER

HARVARD UNIVERSITY, CAMBRIDGE, MASS.

Communicated February 3, 1930

The results which we shall state in this paper have been proved in detail and will be published at an early date.

We are concerned with a calculus of variations problem in the ordinary parametric form, with an integrand which is positive, of class C^4, and homogeneous in the usual way.[1] The domain of the coördinates is a closed region R in m-space.

Let g be an extremal segment in R, and let M_r and M_s be two regular manifolds of class C^3 and of dimensionalities r and s, respectively, $(O \leqq r, s \leqq m - 1)$ which cut g transversally at the distinct points A and B, respectively. Suppose that the integrand F is positively regular along g.

Let K be the class of all regular curves of class D', whose end-points lie on M_r and M_s, respectively. The classical problem of the calculus of variations is to determine necessary and sufficient conditions that the integral take on a proper relative minimum along g from M_r to M_s with respect to its values on curves of the class K. In this paper we present a solution of the problem.

Let H_r be the family of extremals in the neighborhood of g cut transversally by M_r. Let the equations of H_r be given in terms of regular parameters[2] $(t, v_1, \ldots, v_n)$ so chosen that $(v) = (v^\circ)$ gives g. We prove that the equations of H_r can be given in this form with regular parameters.[2] Let $a_{ij}(t)$, $(i, j = 1, \ldots, m)$ be the elements of the jacobian of H_r, with respect to the parameters (t, v), evaluated at $(v) = (v^\circ)$. Points P on g at which the determinant of the elements $a_{ij}(t)$ vanishes are said to be

focal points of M_r, and the order of the vanishing of the determinant is said to be the order of the corresponding focal point. It is proved that the point A is a focal point of M_r of order $n - r$, $n = m - 1$.

Let $t = a$ and $t = b$ $(a < b)$ be the values of the parameter t on g at the points A and B, respectively. Let the focal points of M_r on g exclusive of those on $a < t < b$ be the points: $t = d_1, d_2$, etc., and the points $t = c_1$ c_2, etc., where

$$\ldots c_3 \leqq c_2 \leqq c_1 \leqq a < b \leqq d_1 \leqq d_2 \leqq d_3 \ldots, \text{ etc.} \qquad (1)$$

where focal points are counted according to order so that the equality sign holds between k successive terms of (1) corresponding to a focal point of order k.

Similarly, let the focal points of M_s except those on $a < t < b$ be:

$$\ldots c_3' \leqq c_2' \leqq c_1' \leqq a < b \leqq d_1' \leqq d_2' \leqq d_3' \ldots, \text{ etc.} \qquad (2)$$

THEOREM 1. *In order that the integral take on a relative minimum along g it is necessary that there be no focal points of M_r or of M_s on $a < t < b$, and that:*

$$\ldots c_2' \leqq c_2,\ c_1' \leqq c_1,\ d_1' \leqq d_1,\ d_2' \leqq d_2 \ldots, \text{ etc.} \qquad (3)$$

Let the Weirstrass E-function[3] $E(z, r, \sigma)$ be positive for (z, r) in the neighborhood of $(z, \dot{z})$ on g, and (σ) any set not (o) nor proportional to (r). This will be called the Weirstrass condition.

THEOREM 2. *Suppose that the integrand F is positively regular along g, and that the Weirstrass condition is satisfied. In order that the integral take on a proper relative minimum along g it is sufficient that there be no focal points of M_r or of M_s on $a < t < b$, and that at least one of the following conditions be satisfied:*

$$\ldots, c_2' < c_{n+1},\ c_1' < c_n,\ d_n' < d_1,\ d_{n+1}' < d_2, \ \ldots, \text{ etc.} \qquad (4)$$

The necessary and sufficient conditions of theorems one and two, respectively, do not coincide very closely. In the special case of the problem of the calculus of variations in the plane, conditions (4) are the same as conditions (3), except for the equality signs. This is because in the plane $n = 1$. In this special case, the conditions have been obtained by Bliss.[4]

We have obtained essentially all the conditions as indicated by the following two theorems.

THEOREM 3. *Assume A and B are not conjugate. Let Σ denote any condition on the relative distribution of focal points which is not a consequence of the necessary conditions of Theorem 1. There exists a pair of manifolds M and M' cutting g transversally at A and B, respectively, whose focal points have a distribution which fails to satisfy the condition Σ, and yet for which g gives a weak minimum.*

THEOREM 4. *Let there be no pairs of conjugate points on g and let the Weirstrass and regularity conditions be satisfied. Let* Σ *be any condition on the relative distribution of the points* $d_1, \ldots, d_n$ *and* $d'_1, \ldots, d'_n$ *which is not a consequence of the sufficient conditions (4) of Theorem 2. There exists a pair of manifolds M and M' cutting g transversally at A and B, respectively, whose focal points satisfy the condition* Σ*, and which have no focal points on* $a < t < b$*, and yet for which g fails to give even a weak minimum.*

We have obtained other necessary and sufficient conditions which coincide more closely than the conditions of theorems one and two.

Let H_s be the field of extremals in the neighborhood of g cut transversally by M_s, and let $b_{ij}(\gamma)$, $(i, j = 1, \ldots, m)$ be the elements of the jacobian of H_s with respect to regular parameters $(\gamma, u_1, \ldots, u_n)$ evaluated at $(u) = (u^\circ)$, where (u°) is the set of parameters of g in the field H_s. Let P be a point on g which is not a focal point of M_r or of M_s. Let the parameters (t, v) and (γ, u) be chosen so that at the point P on g:

$$a_{ij} = b_{ij}, (i, j = 1, \ldots, m). \tag{5}$$

THEOREM 5. *The positive and negative type numbers and the nullity of the form D (v), where:*

$$D(v) = a_{ik} F_{r_s r_j} (\dot{a}_{jh} - \dot{b}_{jh}) v_h v_k, (i, j = 1, \ldots, m; k, h = 2, \ldots, m) \tag{6}$$

are numerical invariants of the point P and the manifolds M_r *and* M_s *with respect to admissible*[1] *space transformations and with respect to such admissible parameter transformations*[2] *as preserve the relationship* (5).

The arguments of the partial derivatives of F in (6), are $(z, \dot{z})$ on g at P; and a_{ik}, $\dot{a}_{jh}$, $\dot{b}_{jh}$ are taken at P.

THEOREM 6. *A necessary condition that the integral take on a relative minimum along g is that the form D(v) be positive indefinite, and that there be no focal points of* M_r *or of* M_s *on* $a < t < b$.

THEOREM 7. *Granting that F is positively regular on g and the Weirstrass condition is satisfied, a sufficient condition that the integral take on a proper relative minimum along g is that there be no focal points of* M_r *or of* M_s *on* $a < t < b$*, and the form D(v) be positive definite.*

Thus the problem of the minimum has been solved, but there remains the general problem of classifying extremal segments g cut transversally by M_r and M_s according to the negative type number and the nullity of a fundamental quadratic form.

We cut across the segment of g between M_r and M_s by p successive regular n-manifolds t_i of class C^3 not tangent to g and so close together that no pairs of conjugate points occur on the closed segments of g between successive manifolds. Let R and S be points on M_r and M_s, respectively, and let P_i be a point on t_i. If the points $(R, P_1, \ldots, P_p, S)$ are close to

g they can be joined by unique successive extremal segments forming a broken extremal E. Let (u) be the set of $\mu = r + s + pn$ parameters of which the first r are regular parameters of the point R on M_r, the next n are regular parameters of the point P_1 on t_1, etc., until finally the last s are regular parameters of the point S on M_s. The value of the integral taken along the broken extremal E will be a function of the variables (u) and will be denoted by $J(u)$. The function $J(u)$ will have a critical point when $(u) = (u^\circ)$, where (u°) are the parameters of g in the family of broken extremals E. We define the form Q as

$$Q = J_{ij}u_iu_j, \quad (i, j = 1, \ldots, \mu). \tag{7}$$

where the arguments of the partial derivatives of J are (u°). We classify the extremal segments cut transversally by M_r and M_s according to the negative type number and the nullity of the form Q.

Let P be any point on the open segment of g between M_r and M_s which is not a focal point of M_r or of M_s. Let q_r be the sum of the orders of the focal points of M_r on the open segment of g between M_r and P, and let q_s be the sum of the orders of the focal points of M_s on the open segment of g between P and M_s. Let $D(v)$ be the invariant form (6) constructed at P.

THEOREM 8. *The nullity of the form Q is equal to the nullity of the form $D(v)$.*

THEOREM 9. *The negative type number of the form Q is equal to*

$$q_r + N + q_s \tag{8}$$

where N is the negative type number of the form $D(v)$.

[1] Marston Morse, "The Foundations of the Calculus of Variations in the Large in m-Space" (first paper), *Trans. Amer. Math. Soc.*, **31** (1929), pp. 380–381.

[2] The parameters (t, v) of H_r will be said to be *regular* if the equations are of class C^2, and if $(v) = \text{const.}$ gives the members of H_r as regular 1-manifolds in terms of the parameter t, and if the determinant of $a_{ij}(t)$ vanishes, if at all, at isolated points. A parameter transformation which carries (t, v) into another set of regular parameters is said to be *admissible*.

[3] Bliss, "The Weirstrass E-Function for Problems of the Calculus of Variations in Space," *Trans. Amer. Math. Soc.*, **15** (1914), pp. 369–378.

[4] Bliss, "Jacobi's Criterion When Both End Points Are Variable," *Math. Ann.*, **58** (1903), p. 70.

ON SOME STATISTICAL PROPERTIES OF DOUBLE STARS IN SPACE. I. A FORMULA FOR THE ESTIMATION OF THE PERIOD IN A RELATIVELY FIXED SYSTEM

By Willem J. Luyten

Harvard College Observatory, Cambridge, Massachusetts

Communicated February 15, 1930

The majority of statistics on the frequency of double stars have been based hitherto on the double stars in our catalogs. Hence arise such statements that one star out of nine among those brighter than $6^{m}.5$ is a visual double. When it comes to the actual conditions in space, however, these statistics are as significant as, for example, an analysis of the spectral classes of the stars visible to the unaided eye with respect to the spectral classes of the stars in space. As far as the periods of double stars are concerned, it was Hertzsprung who made the first attempt at estimating the median period of double stars in space.[1] The present writer has shown previously[2] that Hertzsprung's results are incomplete, and rest on arbitrary assumptions. As our knowledge of the stars in the neighborhood of the sun increases, the first defect may be remedied, the second possibly alleviated.

The chief difficulty in the solution of these questions is that of the definition of a double star. At present the known range of physically connected stars is from objects in or near actual contact, to extended systems such as the Ursa Major moving cluster, and ultimately to all open clusters and even globular clusters. Where to draw the limit of a binary system is impossible to say without laying down rules arbitrarily.

It appears to the present writer that at present the line may most conveniently be drawn between such systems as Alpha-Proxima Centauri, and such clusters as the Ursa Major group. In other words, the system α Centauri AB-C has been considered as a binary, but not the system α Canis Majoris-α Coronae Borealis.

The next difficulty which presents itself is that of deciding which period to take in a multiple system. Hertzsprung proposed the rule of always taking the shortest period in any given system, but did not follow it consistently. One circumstance, possibly a result of our incomplete knowledge of double stars, comes to the rescue here, for we do not yet know many systems that approach a real "three-body problem" in the sense that all distances are comparable in size. In all the cases now under consideration we have either a comparatively close binary with a distant companion (α Centauri) or two or more such binaries at comparatively large distances from each other. In such systems it may then be sug-

gested to take as the periods, the period of the close binary on the one hand, as well as that of the distant companion around the close binary, the latter now being treated as a single body.

Having removed the theoretical difficulties by making arbitrary assumptions, there still remains the practical difficulty of estimating the period of a binary star when during the short interval covered by our observations the components remain relatively fixed. In some instances a fair estimate of the period may be derived from the curvature of the path, but this method fails entirely in the case of a distant companion, with a probable period of hundreds of thousands of years. A purely statistical procedure is therefore indicated, and its execution is again facilitated by the fact that for the moment our interest centers mainly on the double stars in the immediate vicinity of the sun, so that the parallaxes may be considered known. Angular distances may thus be converted into linear measure.

According to Kepler's third law we have, for the relation between the period (in years), the major axis (in astronomical units), and the mass (in solar masses):

$$P^2 = a^3/M \tag{1}$$

It appears advantageous to take the logarithm:*

$$\log P = {}^3/_2 \log a - {}^1/_2 \log M \tag{2}$$

Since for double stars in the neighborhood of the sun the luminosities are accurately known, the masses may be estimated from the mass-luminosity curve. The remaining unknown in (2), the semi axis major, presents greater difficulties. Let r be the radius vector in the orbit at the time of observation, d the observed distance at right angles to the line of sight, in astronomical units, e the eccentricity of the orbit and ψ the angle between the radius vector and the line of sight. We may then write:

$$\begin{aligned} r &= a.f(e) \\ d &= r \sin \psi \\ \text{or } d &= a \sin \psi f(e) \end{aligned} \tag{3}$$

and adapting for use in (2):

$$\log a = \log d - \log \sin \psi - \log f(e) \tag{4}$$

whence:

$$\log P = {}^3/_2 \log d - {}^1/_2 \log M - {}^3/_2[\log \sin \psi + \log f(e)] \tag{5}$$

The first two terms on the right hand side are known from observation, and it remains to derive statistical mean values for the other two. There is as yet no reliable evidence for a favored direction in the orientation of double star orbits; we may therefore regard the distributions involved

as random. The mean value of log sin ψ and its standard deviation are then immediately found from the well-known equations

$$\overline{\log \sin \psi} = \int_0^{\frac{\pi}{2}} \log \sin \psi \sin \psi d\psi$$
$$\overline{\log^2 \sin \psi} = \int_0^{\frac{\pi}{2}} \log^2 \sin \psi \sin \psi d\psi \qquad (6)$$

The value of the first integral is $\log 2 - 1 = -0.307$, while the second may be evaluated at 0.2804, giving for the required standard deviation 0.432.

In order to evaluate $\overline{\log f(e)}$ we must bear in mind that all radii vectors are not equally probable; for example, it is much more likely, especially in orbits of great eccentricity, that we catch a binary star in apastron than in periastron. We may therefore take the probability of any radius vector as inversely proportional to the angular focal speed at that instant, and hence the required formula for the mean value of log $f(e)$ and its standard deviation are expressed (for an orbit of semi axis major equal to unity) by:

$$\overline{\log f(e)} = \int_{1-e}^{1+e} \log r . r^2 dr \Big/ \int_{1-e}^{1+e} r^2 dr$$
$$\overline{\log^2 f(e)} = \int_{1-e}^{1+e} \log^2 r . r^2 dr \Big/ \int_{1-e}^{1+e} r^2 dr \qquad (7)$$

The first integral may be solved as:

$$[(1+e)^3 \log (1+e) - (1-e)^3 \log (1-e)]/(6e + 2e^3) - {}^1/_3 \quad (8a)$$

or, expanded into series if desired:

$$(3/2\, e^2 - 1/20\, e^4 - 1/140\, e^6 - 1/504\, e^8 - 1/1320\, e^{10} - \ldots\ldots)/(3 + e^2) \qquad (8b)$$

the general term of which, after the first, is

$$-6/(2n+1).2n(2n-1)(2n-2).e^{2n}$$

Since $u_{2n+2}/u_{2n} = (2n-1)(2n-2)/(2n+3)(2n+2)$, the series is convergent for $|e| < 1$.

The second integral is solved as:

$$(1+e)^3[\{\log(1+e) - {}^1/_3\}^2 + {}^1/_9] - (1-e)^3[\{\log(1-e) - {}^1/_3\}^2 + {}^1/_9]/6e + 2e^3 \qquad (9)$$

The standard deviation in $\log f(e)$ is found by subtracting the square of (8) from (9) and taking the square root. The actual values of $\overline{\log f(e)}$ and its standard deviation are given in the second and third column of table 1.

TABLE 1

e	MEAN	STAND. DEV.	MEAN	STAND. DEV.
0.3	0.044	0.166	0.028	0.101
0.4	0.076	0.212	0.049	0.138
0.5	0.114	0.251	0.074	0.163
0.6	0.159	0.283	0.104	0.184
0.7	0.207	0.307	0.135	0.200
0.8	0.258	0.324	0.168	0.211
0.9	0.309	0.327	0.201	0.213

We may now rewrite (5) in the following form, replacing $\overline{\log f(e)}$ by C:

$$\log P = {}^3/_2 \log d - {}^1/_2 \log M + 0.460 \pm 0.648 - {}^3/_2 C \pm {}^3/_2 \sigma_C \quad (10)$$

or, changing to common logarithms:

$$\log P = {}^3/_2 \log d - {}^1/_2 \log M + 0.200 \pm 0.280 - C' \pm \sigma_{C'} \quad (11)$$

where now the corresponding values of $C' = 0.651\ C$ and its standard deviation may be taken from the fourth and fifth columns of table 1.

Equation (11) may yet be simplified by taking into account the fact that there appears to be a statistical correlation between the period of a binary star and the eccentricity of its orbit. Using the various data that may be derived from current literature, the following rough relationship between P and e may be found. The corresponding values of C' and σ have been entered in the third and fourth columns of table 2, and the total σ, derived by addition into the standard deviation in $\log \sin \psi$, are shown in the fifth column.

TABLE 2

LOG P	e	$C'(e)$	$\sigma_{C'}$	σ_{TOTAL}
3.0	0.62	0.109	0.188	0.337
3.5	0.66	0.122	0.195	0.340
4.0	0.71	0.139	0.202	0.343
4.5	0.76	0.154	0.206	0.345
5.0	0.80	0.168	0.210	0.347
5.5	0.83	0.178	0.212	0.349
6.0	0.85	0.185	0.212	0.351

The first and third columns permit of the establishment of a relation of the sort:

$$C'(e) = 0.028 + 0.027 \log P \quad (12)$$

Substituting this in (11), a procedure not rigidly correct, we obtain:

$$\log P = {}^3/_2 \log d - {}^1/_2 \log M + 0.200 - 0.028 - 0.027 \log P \quad (13)$$

(the third decimal place being rather superfluous), or:

$$\log P = 1.460 \log d - 0.487 \log M + 0.168 \pm 0.35 \qquad (14)$$

P being expressed in years, d in astronomical units and M in solar masses. Since in the calculation of the standard deviation in the last column of table 2 no account has been taken of the fact that there must be a dispersion in eccentricity for a given period, it appears the safest course to adopt the largest value in that column as the standard deviation to be expected in the calculated logarithm of the period. Judging from experience, and taking into consideration possible uncertainties not allowed for in the theory, such as the dispersion on the mass-luminosity curve, uncertainty in the measured separation and the parallax, this value of σ must be regarded as a minimum estimate of the uncertainty in $\log P$. The actual value should be higher, and may well reach 0.40, the actual amount perhaps to be determined from practice.

The significance of formula (14) is that it allows the calculation of the period of a binary system from the angular separation, the parallax and the luminosity, with an uncertainty of 0.40 in the logarithm; that is, the chance is 2 out of 3 that the actual period lies between 0.4 and 2.5 times the computed value.

The usefulness of formula (14) is not limited to cases where the parallax of the star is known. It is clear from its derivation that we may also write:

$$\log a(") = 0.973 \log d(") + 0.112 \pm 0.23. \qquad (15)$$

Although perhaps not strictly correct, this formula is not without statistical usefulness, especially since it already contains an allowance for the increase of eccentricity with length of period (in semi axis major). If, instead of the mean logarithm of a, the actual mean value is desired, it suffices to make the corresponding change in formulae (6) and (7). The evaluation of the resulting integrals is even simpler than that of those used in the present paper, and need not be discussed here.

* Natural logarithms are used as far as and including Equation (10).

[1] *Bull. Astr. Neth.*, **1**, 1922 (149).

[2] *Ann. Harv. Obs.*, **85**, No. 5, 1923 (89).

ON SOME STATISTICAL PROPERTIES OF DOUBLE STARS IN SPACE. II. ON THE MEAN PERIOD OF DOUBLE STARS IN SPACE

BY WILLEM J. LUYTEN

HARVARD COLLEGE OBSERVATORY, CAMBRIDGE, MASSACHUSETTS

Communicated February 15, 1930

In the previous paper a formula was derived for the computation of the period of a binary star when only the parallax, and the angular separation at a given moment are known. Before actually using such a formula, it is always well to test it on a group of stars for which the period, to be estimated from the formula, is already known with reasonable accuracy. For the data required in this computation, the writer is under obligation to Dr. W. H. Van den Bos who kindly supplied the results of the latest known measures of these double stars—in many cases his own unpublished measures.

Table 1 contains the data needed for the actual comparison for 15 binary stars of known parallax and period, arranged in order of increasing period, and designated by the name under which these stars are known to double-star observers. The second, third, and fourth columns contain the angular separation, according to the latest known measure, the adopted parallax, and the mass of the system. The fifth column gives the logarithm of the period, calculated by means of the formula derived in the previous paper, which formula is again repeated at the foot of table 1. The sixth and seventh columns show the period derived by double star calculators, and its logarithm, while the eighth column gives the differences O-C.

According to formula (1) the standard deviation to be expected in any individual logarithm of period is 0.34 (as a minimum), hence that in the mean of 15 periods should be 0.09. From table 1 we notice that the mean value of log P as calculated from the formula is 1.83 ± 0.09, while that of the observed periods is 1.89. The difference O-C = +0.06, only two-thirds of the standard deviation, is small enough to inspire confidence in formula (1). It may well be explained by accidental errors; yet the possibility should not be excluded that formula (1) requires a small systematic correction, resulting from the effect of causes not allowed for in its derivation, such as observational selection and discovery chance. The influence of the latter, however, would be expected to act in the opposite sense. From the internal discordances of the values of O-C in the last column, a standard deviation of 0.27 would be indicated for an individual estimate made by means of the formula. Since this value is smaller than

that derived from the theory, it again inspires confidence in the formula, and we may assume that its present small size is accidental.

Formula (1) may now be judged to be reasonably reliable, and we may thus proceed with its application to the problem of the estimation of the mean period of double stars in space. For the definitions used in judging what is a double star, and what is, or are the periods in a multiple system, reference must be made to the previous paper. It appears to the writer

TABLE 1

STAR	d	π	MASS	LOG P	$P(0)$	LOG $P(0)$	O-C
β 733	0".72	0".100	1.0	1.43	26.7	1.43	0.00
ζ Her (Σ2084)	1.2	0.112	2.1	1.51	34.5	1.54	+0.03
α CMi	3±	0.307	1.6	1.54	40.2	1.60	+0.06
Mlb 4 AB	1.26	0.140	1.1	1.40	42.2	1.62	+0.22
μ Her BC	0.74	0.107	0.90	1.43	43.0	1.63	+0.20
Krü 60 AB	1.36	0.258	0.45	1.06	44.3	1.65	+0.59
α CMa AB	10.6	0.366	3.4	2.07	50.0	1.70	−0.37
ξ UMa AB-CD	1.61	0.145	1.4	1.63	59.9	1.78	+0.15
α Cen AB	7.02	0.757	2.04	1.44	80.1	1.90	+0.46
70 Oph (Σ 2272)	6.56	0.194	1.7	2.32	87.7	1.94	−0.38
Brsb 13	3.60	0.140	0.8	2.32	130?*	2.11	−0.21
ξ Boo	3.01	0.173	1.0	2.01	151	2.18	+0.17
Δ5 (p Eri)	9.22	0.165	2.32	2.58	219	2.34	−0.22
o_2 Eri BC	4.52	0.200	0.64	2.27	248	2.39	+0.12
η Cas	8.14	0.180	1.4	2.56	508	2.71	+0.15
				1.83		1.89	+0.06
				±0.09			±0.07

* According to the orbit derived by Van den Bos (B. A. N., **2,** 29, and **3,** 154) the period of *Brsb* 13 = 41*G Arae* is 100.9 years. From his own recent measures, however, Dr. Van den Bos concludes that this period is much too short, and estimates the real period at something like 130 years.

$$\text{Log } P = 1.460 \log d - 0.487 \log M + 0.168 \pm 0.34 \qquad (1)$$

that the solution of the problem may best be attempted by calculating the periods of all double stars nearer than a given distance—ten parsecs in the present investigation. The material has been taken from the list of stars in Harvard Annals, Vol. **85,** No. 5. To these have been added δ *Trianguli,* since the re-determination of its parallax at the McCormick Observatory now leaves little doubt about its being nearer than ten parsecs. Among southern stars, the new Yale parallax of *p Eridani* places this binary well within our limits, while, for the sake of completeness, ζ *Reticuli* has also been included, although the parallax of this star is known with but scant measure of certainty. It is possible that another system of parallaxes than the one used here would place OΣ547, Σ 1280, Brsb 5 and β 733 = 85 *Pegasi* further away than 10 parsecs; while it is almost certain that OΣ539 AC, and Chri 2448 will fall outside those

limits. In an investigation such as the present, the inclusion of these stars does no harm, provided that no other double star has been omitted nearer than the furthest of those in the present list. This is believed and hoped to be the case, insofar as contemporary parallax information

TABLE 2

STAR	PERIOD	LOG P
δ Tri	0.0272	−1.57
ξ UMa CD	0.027 ±	−1.56
χ Dra	0.769	−0.11
η Boo	1.36	+0.13
ξ UMa AB	1.82	+0.26

TABLE 3

NAME	d	π	d(A. U.)	ABS. MAG.	MASS	LOG P
OΣ 547 AB	4".8	0".100	48	9.3, 9.4	0.77	2.68
OΣ 547 AB-C	330	0.100	3300	9.3, 9.4, 10.2	1.11	5.28
Bo 187	39	0.281	139	10.4, 13.0	0.51	3.44
Millb. 377	2.3	0.115	20	10. 13.	0.51	2.31
κ Tuc AB (h 3423)	5.1	0.11?	46 ±	5.2, 7.4	1.50	2.51
Lac 353 CD* (I 27)	1.2	0.11?	11 ±	7.8, 8.8	1.03	1.58
κ Tuc-Lac 353 AB-CD	319.4	0.11?	2875 ±	5.2, 7.4, 7.8, 8.8	2.53	5.02
Lpz II 961	150	0.142	1060	6.6, 12.2	0.91	4.60
ζ Ret	310	0.1?	3100 ±	5.2, 5.5	1.8	5.13
o_2 Eri A-BC†	83	0.200	415	6.0 + mass 0.64	1.38	3.92
γ Lep (HV 50)	95	0.149	639	4.7, 7.3	1.58	4.17
Σ 1280	5.2	0.100	52	8.7, 9.1	0.86	2.70
Σ 1321	19.4	0.163	119	9.1, 9.1	0.84	3.24
Hu 1128	5.0	0.104	48	5.6, 14	0.99	2.62
OΣ 539 AC	7.8	0.095	87	7.3, 9.7	0.95	3.01
Brs 5 (P = 342^y??)	1.57	0.100	15.7	8.0, 8.0	1.00	1.91
α Cen AB-C††	6740	0.760	8860	15.5 + mass 2.04	2.20	5.77
Sh 190	14.2	0.181	78.8	7.1, 10.2	0.94	2.95
Chri 2448	63.7	0.095	670	8.8 10.5	0.76	4.35
W-Ott 5811	512	0.160	3210	10.2, 12.4	0.56	5.41
Sh 243 AB	4.3	0.174	24.8	6.5, 6.5	1.36	2.14
36 Oph-30 Sco AB-C	730	0.174	4200	6.5, 6.5, 7.8	1.89	5.32
Mlb 4 AB-C	31	0.144	215	11.1 + mass 1.1	1.40	3.51
μ Her A-BC	32	0.105	305	3.7 + mass 0.9	2.20	3.64
Σ 2398	16	0.294	54.4	11.1, 11.7	0.55	2.83
h 5173	9	0.243	37	7.2, 13.2	0.78	2.51
61 Cyg (Σ 2758)	21	0.300	70	8.0, 8.7	0.95	2.87

* The period of I 27 may well be 100 ± years.

† A period of about 8600 years is estimated from circular orbit.

†† The difference in parallax has not been taken into account since this is statistically accounted for in formula (1).

goes. The next applicants for the list appear to be γ *Virginis*, B.D. + 35° 2436, Σ 1819, and β 996. Concerning one star definitely within the limits there exists some uncertainty as to its duplicity, viz., π_3 *Orionis*.*

The actual material at our disposal consists of the 15 stars listed in table 1, 5 spectroscopic binaries, about which information is given in table 2, and 27 other doubles for which the data are shown in table 3—arranged in the same manner as the first five columns of table 1, except that now two extra columns have been inserted, giving the conversion of the observed angular separation into astronomical units and the absolute magnitudes of the components, which have now served for the calculation of the mass. The stars have been arranged in order of R. A. and are designated by their usual double-star names, but they may be easily identified with those in *H. A.*, Vol. **85,** No. 5.

Using the fifteen logarithms of periods from the seventh column of table 1, the five from table 2 and the twenty seven from the seventh column of table 3 we now have 47 values of log P for a distribution. These 47 values are given in table 4, arranged in order of numerical size, while the statistical description of their distribution is shown below.

TABLE 4

−1.57	1.54	1.70	2.14	2.51	2.87	3.64	5.13
−1.56	1.58	1.78	2.18	2.62	2.95	3.92	5.28
−0.11	1.60	1.90	2.31	2.68	3.01	4.17	5.32
+0.13	1.62	1.91	2.34	2.70	3.24	4.35	5.41
+0.26	1.63	1.94	2.39	2.71	3.44	4.60	5.77
+1.43	1.65	2.11	2.51	2.83	3.51	5.02	

Median 2.51; quartiles at 1.65 and 3.54 indicating a standard deviation of 1.41. Arithmetic mean 2.57 ± 0.23, standard deviation, calculated from the actual squares, 1.59; after allowing for dispersion in estimated values of log P from formula (1), $\sigma = 1.57 \pm 0.16$.

The distribution is irregular, which is not surprising since there are only 47 items. The curious part, however, is that most of the unexpected peak between log $P = 1.5$ and 2.0 is due to actually determined periods from table 1, no less than five out of fifteen lying between 40 and 50 years. Under these conditions it should be considered as accidental that the agreement between the median and the mean is so close, although this does inspire some confidence in the statistical stability of the result.

It was remarked before in *Harvard Annals*, **85,** 1923 (90) that the median value of log P for double stars in space was probably not far from 2.5 to 2.8. The lower limit of these two has now been reached by the observations, and it remains to be seen to what extent these observations may be considered as unselectively representative of the conditions in space.

Examination of the list of stars in *H. A.*, Vol. **85,** and comparison with the present list of binaries calls forth the following comments:

(*a*) The known spectroscopic binaries are all brighter than the sixth apparent magnitude. Whether or not spectroscopic binaries or eclipsing

binaries (such as α Geminorum C) exist among the fainter stars is at present not capable of determination by observation.

(*b*) For the visual binaries the search may be taken as complete down to the limits of the double star Durchmusterung, say about 9.0 on the Harvard scale. Below this limit nothing is known with certainty, and the only star fainter than this limit in the present list is B. D. $+ 66° 34$.

(*c*) As far as distant companions are concerned, our present knowledge is based entirely upon accidental discoveries (with the exception of Proxima Centauri, which was deliberately searched for by Innes).

In view of these facts it is almost impossible to say what the probable conditions in space are, and thus, in the absence of any reliable information, we may sum up our ignorance by stating that the chances are as good that the unknown binaries are of longer periods than that they are of shorter periods than the mean derived here. Likewise, it appears reasonable that the actual standard deviation is not much larger than the one derived here.

The final conclusion concerning the periods of binary stars in space may therefore be stated as follows: the median-mean logarithm of the periods (corresponding to the median-geometric mean of the actual periods) is probably in the neighborhood of 2.5, corresponding to a little more than 300 years, while the dispersion in these logarithms is probably not more than 1.70, in other words, half the binaries in space may be expected to have periods between the limits of 20 and 4000 years.

* Cf. *Astrophys. Journ.*, **48**, 1918 (270).